

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: January 30, 2001, 16:45:21 ; Search time 183.67 seconds
(without alignments)
11.356 Million cell updates/sec

Title: US-09-518-931-2_COPY_57_117

Perfect score: 371

Sequence: 1 FVQPCRRDPTTCGPCPPR.....REEEARACHATHNRACRRT 61

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database : A_Geneseq_36.*

- 1: /SIDSL1/gcgdata/geneseq/geneseq/AA1980.DAT.*
- 2: /SIDSL1/gcgdata/geneseq/geneseq/AA1981.DAT.*
- 3: /SIDSL1/gcgdata/geneseq/geneseq/AA1982.DAT.*
- 4: /SIDSL1/gcgdata/geneseq/geneseq/AA1983.DAT.*
- 5: /SIDSL1/gcgdata/geneseq/geneseq/AA1984.DAT.*
- 6: /SIDSL1/gcgdata/geneseq/geneseq/AA1985.DAT.*
- 7: /SIDSL1/gcgdata/geneseq/geneseq/AA1986.DAT.*
- 8: /SIDSL1/gcgdata/geneseq/geneseq/AA1987.DAT.*
- 9: /SIDSL1/gcgdata/geneseq/geneseq/AA1988.DAT.*
- 10: /SIDSL1/gcgdata/geneseq/geneseq/AA1989.DAT.*
- 11: /SIDSL1/gcgdata/geneseq/geneseq/AA1990.DAT.*
- 12: /SIDSL1/gcgdata/geneseq/geneseq/AA1991.DAT.*
- 13: /SIDSL1/gcgdata/geneseq/geneseq/AA1992.DAT.*
- 14: /SIDSL1/gcgdata/geneseq/geneseq/AA1993.DAT.*
- 15: /SIDSL1/gcgdata/geneseq/geneseq/AA1994.DAT.*
- 16: /SIDSL1/gcgdata/geneseq/geneseq/AA1995.DAT.*
- 17: /SIDSL1/gcgdata/geneseq/geneseq/AA1996.DAT.*
- 18: /SIDSL1/gcgdata/geneseq/geneseq/AA1997.DAT.*
- 19: /SIDSL1/gcgdata/geneseq/geneseq/AA1998.DAT.*
- 20: /SIDSL1/gcgdata/geneseq/geneseq/AA1999.DAT.*
- 21: /SIDSL1/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	371	100.0	170	19 W63623	Human tumour necro
2	371	100.0	245	20 Y28449	A human tumour nec
3	371	100.0	271	20 Y42184	Human mFLINT #1 pr
4	371	100.0	300	19 W66102	Amino acid sequenc
5	371	100.0	300	19 W63622	Human tumour necro
6	371	100.0	300	20 Y03099	Human lung TNF-rec
7	371	100.0	300	20 Y42182	Human FLINT #1 pro
8	371	100.0	300	20 Y17479	Mammalian tumour n
9	371	100.0	300	20 Y06817	Human DCR3 polypep
10	371	100.0	300	20 W97749	Human tumour necro
11	371	100.0	300	20 W95082	Orphan receptor (H
12	366	98.7	273	20 Y42185	Human mFLINT #2 pr

13	366	98.7	302	20 Y42183	Human FLINT #2 pro
14	357	96.2	300	21 Y77458	Human TNF receptor
15	193	52.0	153	20 Y22222	Human TNFR superfa
16	189	50.9	215	20 W93585	Human hAPO6 protei
17	150	40.4	360	17 R99938	Mutated OCIF, OCIF
18	148	39.9	401	17 R99933	Mutated OCIF, OCIF
19	147	39.6	143	17 R99946	Mutated OCIF, OCIF
20	147	39.6	145	17 R99930	Osteoclastogenesis
21	147	39.6	154	17 R99929	Osteoclastogenesis
22	147	39.6	187	17 R99950	Mutated OCIF, OCIF
23	147	39.6	187	21 Y77464	Primate protein se
24	147	39.6	197	17 R99945	Mutated OCIF, OCIF
25	147	39.6	272	17 R99944	Mutated OCIF, OCIF
26	147	39.6	321	17 R99949	Mutated OCIF, OCIF
27	147	39.6	326	17 R99940	Mutated OCIF, OCIF
28	147	39.6	327	17 R99941	Mutated OCIF, OCIF
29	147	39.6	349	20 W83928	Human FTHMA-070 pa
30	147	39.6	351	17 R99943	Mutated OCIF, OCIF
31	147	39.6	359	17 R99939	Mutated OCIF, OCIF
32	147	39.6	380	17 R99924	Mature osteoclasto
33	147	39.6	390	17 R99357	Human tumour necro
34	147	39.6	391	19 W53238	Human OCIF genome
35	147	39.6	393	17 R99948	Mutated OCIF, OCIF
36	147	39.6	395	19 W57636	Modified TRL recep
37	147	39.6	399	17 R99942	Mutated OCIF, OCIF
38	147	39.6	401	17 R99932	Mutated OCIF, OCIF
39	147	39.6	401	17 R99934	Mutated OCIF, OCIF
40	147	39.6	401	17 R99925	Full length osteoc
41	147	39.6	401	17 R99931	Mutated OCIF, OCIF
42	147	39.6	401	18 W38345	Human osteoprotege
43	147	39.6	401	19 W57635	TRL receptor prote
44	147	39.6	401	19 W53239	Human OCIF genome
45	147	39.6	401	20 W83926	Human FTHMA-070 pr

ALIGNMENTS

RESULT 1	
W63623	
ID W63623 standard; Protein: 170 AA.	
XX	
AC W63623;	
XX	
DT 26-OCF-1998 (first entry)	
XX	
DE Human tumour necrosis factor receptor-6 beta protein.	
XX	
KW Human tumour necrosis factor receptor-6 alpha; TNFR-6 beta;	
KW endothelial cells; keratinocytes; normal prostate; apoptosis;	
KW prostate tumour tissue.	
XX	
OS Homo sapiens.	
XX	
FH Key	Location/Qualifiers
FT Peptide	1..30
FT Protein	31..170
FT Region	note="TNFR-6 beta"
FT	31..166
FT	/note= "Soluble extracellular domain"
XX	
PN W09830694-A2.	
XX	
PD 16-JUL-1998.	
XX	
PF 13-JAN-1998; 98WO-US00153.	
XX	
PR 14-JAN-1997; 97US-0035496.	
XX	
PA (HUMA-) HUMAN GENOME SCI INC.	
XX	
PI Ebner R, Feng P, Gentz RL, Ni J, Ruben SM, Yu G;	
XX	

DR WPI; 1998-399142/34.
 DR N-PSDB; V39086.
 XX Human tumour necrosis factor receptors 6-alpha and 6-beta - used in
 PT the diagnosis of immune system-related disorder(s)
 XX
 XX Claim 20; Fig 2A-2B; 91pp; English.
 XX
 CC The present sequence represents the human tumour necrosis factor
 CC receptor-6 beta (TNFR-6 beta) protein. The invention also provides
 CC for the TNFR-6 alpha protein (#63622). TNFR-6 alpha and TNFR-6 beta are
 CC members of the tumour necrosis factor receptor (TNFR) family. TNFRs
 CC are expressed in endothelial cells, keratinocytes, normal prostate and
 CC prostatic tumour tissue. For a number of disorders of these cells,
 CC particularly of the immune system, substantially altered (whether
 CC increased or decreased) levels of TNFR-6 alpha and/or TNFR-6 beta gene
 CC expression can be detected, therefore the TNFR-6 alpha and TNFR-6 beta
 CC polypeptides, nucleic acids and antibodies are claimed to be useful in
 CC the diagnosis of such disorders. Mutations of the TNFR-6 alpha and
 CC TNFR-6 beta genes can also be detected. The TNFR polypeptides are
 CC also claimed to be useful for identifying ligands which may be useful
 CC in the treatment of apoptosis related disorders.
 XX
 XX Sequence 170 AA;
 SQ

Query Match 100.0%; Score 371; DB 19; Length 170;
 Best Local Similarity 100.0%; Pred. No. 3.2e-30;
 Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVORCRRDSPTTCGPPRRHYTFQWNYLERCRYCNVLCGEREEARACHATHNRCRCR 60
 DB 57 fvdprcrdspttcgpprrhytqfwnylercrycnvlgereearachathnrcrcr 116
 QY 61 T 61
 DB 117 t 117

RESULT 2
 Y28449
 ID Y28449 standard; Protein; 245 AA.
 AC Y28449;
 XX
 XX 29-SEP-1999 (first entry)
 XX
 XX A human tumour necrosis factor-R2-like proteins (TR2P)-1.
 XX Human tumour necrosis factor-R2-like protein; TR2P; achondroplasia;
 KW osteoporosis; developmental disorder; Cushing's syndrome;
 KW muscular dystrophy; epilepsy; hereditary neuropathy;
 KW Charcot-Marie-Tooth disease; neurofibromatosis; hypothyroidism;
 KW hydrocephalus; seizure disorder; cerebral palsy; spinal bifida;
 KW congenital glaucoma; cataract; sensorineural hearing loss;
 KW reproductive disorder; infertility; ovulatory defect; endometriosis;
 KW autoimmune disorder; ectopic pregnancy; teratogenesis; spermatogenesis;
 KW immunological disorder; AIDS; Addison's disease; allergy; bronchitis;
 KW atherosclerosis; diabetes mellitus; Chron's disease; lupus;
 KW irritable bowel syndrome; multiple sclerosis; infection;
 KW neoplastic disorder; adenocarcinoma; leukaemia; lymphoma; melanoma;
 KW myeloma; sarcoma.
 XX
 XX Homo sapiens.
 XX
 XX WO9931128-A2.
 XX
 XX 24-JUN-1999.
 XX
 XX 02-DEC-1998; 98WO-US25649.
 XX
 XX 16-DEC-1997; 97US-0991945.
 XX

(INCY-) INCYTE PHARM INC.
 Au-Young J, Bandman O, Hillman JL, Kaser MR, Tang YT;
 WPI; 1999-457916/38.
 DR N-PSDB; X89503.
 XX
 PT New tumour necrosis factor-R2-like protein - useful in the treatment
 PT of osteogenesis, developmental, reproductive, immunological and
 XX neoplastic disorders
 PS Claim 1; Fig 1A-C; 81pp; English.
 XX
 CC The present sequence represents a human tumour necrosis factor-R2-like
 CC protein (TR2P). The protein is used to treat and prevent osteogenesis,
 CC developmental, reproductive, immunological and neoplastic disorders, and
 CC also to diagnose disorders associated with TR2 protein expression. Such
 CC disorders include osteogenesis disorders such as achondroplasia and
 CC osteoporosis, developmental disorders such as Cushing's syndrome,
 CC muscular dystrophy, and epilepsy, hereditary neuropathies such as
 CC Charcot-Marie-Tooth disease, and neurofibromatosis, hypothyroidism,
 CC hydrocephalus, seizure disorders such as cerebral palsy and spinal
 CC bifida, congenital glaucoma, cataract, or sensorineural hearing loss,
 CC reproductive disorders such as infertility, ovulatory defects and
 CC endometriosis, autoimmune disorders, ectopic pregnancy and teratogenesis,
 CC disruption of spermatogenesis, immunological disorders such as AIDS,
 CC Addison's disease, allergies, bronchitis, atherosclerosis, diabetes
 CC mellitus, Chron's disease, lupus and irritable bowel syndrome, multiple
 CC sclerosis, viral, fungal, helminthic, parasitic and protozoal infections,
 CC and neoplastic disorders including adenocarcinoma, leukaemia, lymphoma,
 XX melanoma, myeloma, sarcoma, and teratocarcinoma.
 XX Sequence 245 AA;
 SQ

Query Match 100.0%; Score 371; DB 20; Length 245;
 Best Local Similarity 100.0%; Pred. No. 4.5e-30;
 Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVORCRRDSPTTCGPPRRHYTFQWNYLERCRYCNVLCGEREEARACHATHNRCRCR 60
 DB 57 fvdprcrdspttcgpprrhytqfwnylercrycnvlgereearachathnrcrcr 116
 QY 61 T 61
 DB 117 t 117

RESULT 3
 Y42184
 ID Y42184 standard; Protein; 271 AA.
 XX
 AC Y42184;
 XX
 XX 17-DEC-1999 (first entry)
 XX
 XX Human mFLINT #1 protein sequence.
 XX
 XX Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
 KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
 KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
 KW reperfusion-associated injury; aplastic anaemia; differentiation;
 KW growth; myelodysplastic syndrome; pancytopenic condition;
 KW myocardial ischaemia.
 XX
 XX Homo sapiens.
 XX
 XX WO9950413-A2.
 XX
 XX 07-OCT-1999.
 XX
 XX 30-MAR-1999; 99WO-US06797.
 XX

PR 30-MAR-1998; 98US-0079856.
 PR 20-MAY-1998; 98US-0086074.
 PR 09-SEP-1998; 98US-0099643.
 PR 17-DEC-1998; 98US-0112577.
 PR 18-DEC-1998; 98US-0112703.
 PR 18-DEC-1998; 98US-0112933.
 PR 22-DEC-1998; 98US-0113407.
 XX (ELIL) LILLY & CO ELI.
 XX
 PI Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
 PI Hul KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
 PI Song HY, Wang J, Wu X, Zuckerman SH;
 XX
 DR WPI: 1999-591319/50.
 DR N-PSDB: 225377.
 XX
 XX Use of mature FLINT for treating acute liver failure, inflammation,
 PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
 PT and proinflammatory activity
 XX
 PS Claim 31; Fig 3; 99pp; English.
 XX
 CC The present invention describes therapeutic applications of mature FLINT
 CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
 CC (mFLINT), which is a member of the tumour necrosis factor receptor
 CC superfamily, is used for treating acute liver failure, inflammation of
 CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
 CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
 CC injury or disorder such as hypercoagulation (including use with
 CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
 CC or disorder, Type I diabetes, cancer, cell damage or damage to an
 CC innocent bystander tissue that is induced by a chemotherapeutic agent or
 CC therapeutic irradiation, treating haematopoietic progenitor cells that
 CC have been exposed to therapeutic radiation or chemotherapy, aplastic
 CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
 CC also used for promoting the growth or differentiation of a haematopoietic
 CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
 CC resulting from abnormal myocardial ischaemia. The present sequence
 CC represents human mFLINT.
 XX
 SQ Sequence 271 AA;
 Query Match 100.0%; Score 371; DB 20; Length 271;
 Best Local Similarity 100.0%; Pred. No. 5e-30;
 Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 FVQPCRRDSTTCGCPPRHYTFWNYLCRYCNVLCGEREEARACHATHNRACR 60
 Db 28 fvqpcrrdspttcgpcprhytqfwnylcrvcnvlcgereearachathnrcrcr 87
 Qy 61 T 61
 Db 88 t 88
 RESULT 4
 W66102
 ID W66102 standard; Protein; 300 AA.
 XX
 AC W66102;
 XX
 DT 02-DEC-1998 (first entry)
 DE Amino acid sequence of tumour necrosis related receptor (TR4).
 XX
 KW Human; tumour necrosis related receptor; TR4; agonist; antagonist;
 KW inhibition; chronic; acute; inflammation; arthritis; septicemia;
 KW autoimmune disease; transplant rejection; stroke; cancer;
 KW Alzheimer's disease.
 XX
 OS Homo sapiens.

XX EP861850-A1.
 PN
 XX
 PD 02-SEP-1998.
 XX
 PF 20-JAN-1998; 98EP-0300382.
 XX
 PR 04-FEB-1997; 97US-0794796.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX
 PI Emery J, Tan KB, Truneh A, Young PR;
 XX
 DR WPI: 1998-508248/44.
 DR N-PSDB: V07654.
 XX
 XX New DNA encoding tumour necrosis related receptor - used to treat
 PT and prevent e.g. inflammation, arthritis, septicemia, autoimmune
 PT diseases, transplant rejection, infection, stroke, ischaemia, ARDS,
 PT restenosis, AIDS, bone disorders and cancer
 XX
 PS Claim 1; Fig 1; 2lpp; English.
 XX
 CC This is the amino acid sequence of the human tumour necrosis related
 CC receptor (TR4), used in the method of the invention. The TR4 protein
 CC or its agonist can be used to treat a subject in need of enhanced
 CC TR4 polypeptide activity. The antagonist is used to inhibit TR4
 CC polypeptide activity. The active agents can be used for the
 CC treatment and prevention of diseases such as chronic and acute
 CC inflammation, arthritis, septicemia, autoimmune diseases, transplant
 CC rejection, stroke, cancer, Alzheimer's disease.
 XX
 SQ Sequence 300 AA;
 Query Match 100.0%; Score 371; DB 19; Length 300;
 Best Local Similarity 100.0%; Pred. No. 5.5e-30;
 Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 FVQPCRRDSTTCGCPPRHYTFWNYLCRYCNVLCGEREEARACHATHNRACR 60
 Db 57 fvqpcrrdspttcgpcprhytqfwnylcrvcnvlcgereearachathnrcrcr 116
 Qy 61 T 61
 Db 117 t 117
 RESULT 5
 W63622
 ID W63622 standard; Protein; 300 AA.
 XX
 AC W63622;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Human tumour necrosis factor receptor-6 alpha protein.
 XX
 KW Human tumour necrosis factor receptor-6 alpha; TNFR-6 alpha; TNFR-6 beta;
 KW endothelial cells; keratinocytes; normal prostate; apoptosis;
 KW prostate tumour tissue.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..30
 FT Protein 31..300
 FT /note= "TNFR-6 alpha"
 FT Region 31..282
 FT /note= "Soluble extracellular domain"
 XX
 PN W09830694-A2.
 XX

```

PD 16-JUL-1998.
XX
PF 13-JAN-1998; 98WO-US00153.
XX
PR 14-JAN-1997; 97US-0035496.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ebner R, Feng P, Gentz RL, Ni J, Ruben SM, Yu G;
XX
DR WPI: 1998-399142/34.
DR N-PSDB; V39085.
XX
PT Human tumour necrosis factor receptors 6-alpha and 6-beta - used in
PT the diagnosis of immune system-related disorder(s)
XX
PS Claim 20; Fig 1; 91pp; English.
XX
CC The present sequence represents the human tumour necrosis factor
CC receptor-6 alpha (TNFR-6 alpha) protein. The invention also provides
CC for the TNFR-6 beta protein (W63623). TNFR-6 alpha and TNFR-6 beta are
CC members of the tumour necrosis factor receptor (TNFR) family. TNFRs
CC are expressed in endothelial cells, keratinocytes, normal prostate and
CC prostate tumour tissue. For a number of disorders of these cells,
CC particularly of the immune system, substantially altered (whether
CC increased or decreased) levels of TNFR-6 alpha and/or TNFR-6 beta gene
CC expression can be detected, therefore the TNFR-6 alpha and TNFR-6 beta
CC polypeptides, nucleic acids and antibodies are claimed to be useful in
CC the diagnosis of such disorders. Mutations of the TNFR-6 alpha and
CC TNFR-6 beta genes can also be detected. The TNFR polypeptides are
CC also claimed to be useful for identifying ligands which may be useful
CC in the treatment of apoptosis related disorders.
XX
SQ Sequence 300 AA;

Query Match 100.0%; Score 371; DB 19; Length 300;
Best Local Similarity 100.0%; Pred. No. 5.5e-30;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVORPCRRDPTTCGPPRRHYTFQWNYLCRYCNVLCGEREEARACHATHNRACRCR 60
DB 57 fvqpcrrdspttcgpprrhytqfwnylcrvcnvlcgereearachathnrcr 116
QY 61 T 61
DB 117 t 117

RESULT 6
Y03099
ID Y03099 standard; Protein; 300 AA.
XX
AC Y03099;
XX
DT 09-DEC-1999 (first entry)
XX
DE Human lung TNF-receptor protein.
XX
KW Tumour necrosis factor; TNF; TNF receptor; human; lung; gene therapy;
KW detection; immunoassay; diagnosis; disease; immune system; tumour;
KW osteogenic system; cardiovascular system; central nervous system; asthma;
KW peripheral nervous systems; transplant incompatibility; antitumor;
KW rheumatoid arthritis; antiasthmatic; antiarthritic.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 134..1036
FT /*tag= a
FT /product= "TNF-receptor"
XX
PN DE19809978-A1.

XX 16-SEP-1999.
XX
XX 09-MAR-1998; 98DE-1009978.
XX
XX 09-MAR-1998; 98DE-1009978.
XX
XX (BADI ) BASF AG.
XX
XX Kroeger B;
XX
XX WPI: 1999-519473/44.
XX N-PSDB; Z09998.
XX
XX New soluble member of tumor necrosis factor receptor family, useful for
XX identification specific modulators and for treating disease e.g. tumors
XX
XX Claim 1; Page 8-9; 10pp; German.
XX
XX This invention describes a novel tumour necrosis factor (TNF) receptor
XX (I) isolated from human lung tissue. (I) is used: (i) to raise specific
XX antibodies (Ab); (ii) to screen for specific (ant)agonists or ligands
XX (A), potential therapeutic agents; and (iii) therapeutically (optionally
XX expressed from a gene therapy vector) in conditions associated with a
XX deficit of (I). Ab are used: (a) for qualitative or quantitative
XX detection of (I) in standard immunoassays (for diagnosis of disease, or
XX susceptibility, or for monitoring); and (b) as therapeutic inhibitors in
XX cases where (I) is overexpressed. Nucleic acid (II) that encodes (I) is
XX used: (A) for recombinant production of (I); (B) also its oligonucleotide
XX fragments, in standard hybridization and/or amplification assays; (C) as
XX source of antisense molecules or ribozymes; and (D) to produce transgenic
XX animals (for studying (patho)physiology of (I)). Diseases possibly
XX associated with under- or over-expression of (I) are those of the immune,
XX osteogenic, cardiovascular and central or peripheral nervous systems, The
XX tumors, transplant incompatibility, asthma and rheumatoid arthritis. The
XX products of the invention have antitumor, antiasthmatic and
XX antiarthritic activity. This sequence represents the TNF-receptor of the
XX invention.
XX
SQ Sequence 300 AA;

Query Match 100.0%; Score 371; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 5.5e-30;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVORPCRRDPTTCGPPRRHYTFQWNYLCRYCNVLCGEREEARACHATHNRACRCR 60
DB 57 fvqpcrrdspttcgpprrhytqfwnylcrvcnvlcgereearachathnrcr 116
QY 61 T 61
DB 117 t 117

RESULT 7
Y42182
ID Y42182 standard; Protein; 300 AA.
XX
AC Y42182;
XX
DT 17-DEC-1999 (first entry)
XX
DE Human FLINT #1 protein sequence.
XX
KW Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
KW reperfusion-associated injury; aplastic anaemia; differentiation;
KW growth; myelodysplastic syndrome; pancytopenic condition;
KW myocardial ischaemia.
XX

```



```

XX PN WO9914330-A1.
XX FT 25-MAR-1999.
XX PD
XX PF 18-SEP-1998; 98WO-US19661.
XX PR 30-JUL-1998; 98US-0094640.
XX PR 18-SEP-1997; 97US-0059288.
XX PA (GETH ) GENENTECH INC.
XX PI Ashkenazi AJ, Botstein D, Dodge KH, Goddard A, Gurney AL;
XX PI Kim KJ, Lawrence DA, Pitti R, Roy MA, Tumas DB;
XX PI Wood WI;
XX DR WPI; 1999-244032/20.
XX DR N-PSDB; X32744.
XX PT Dcr3 polypeptide related to tumor necrosis factor receptor
XX PS Claim 5; Fig 1; 88pp; English.
XX XX This represents a human Dcr3 polypeptide, a homologue of tumour necrosis
XX CC factor receptor (TNFR) polypeptide. Host cells containing a vector
XX CC comprising the Dcr3 nucleic acid can be used for the recombinant
XX CC expression of the protein. Dcr3 binds to Fas ligand, so it (or its
XX CC chimeras) are useful for modulating apoptosis in mammalian cells, also
XX CC other Fas-ligand induced activities, particularly to inhibit T cell
XX CC mediated immune responses, e.g. in treatment of allergy, asthma,
XX CC rheumatoid arthritis, Crohn's disease, guest vs. host disease etc. Dcr3
XX CC may also be used to identify specific binding proteins, potential
XX CC inhibitors. Antibodies against Dcr3 are used to treat cancer,
XX CC specifically of the lung and colon, also in diagnosis and for affinity
XX CC purification of the protein. Detecting mutations in the gene for Dcr3 is
XX CC also used to diagnose cancer, or predisposition to it. Dcr3 nucleic acid
XX CC is useful as hybridization probe to detect genomic or related sequences;
XX CC for chromosome and gene mapping; as source of antisense sequences; for
XX CC expression of recombinant Dcr3 and to generate transgenic animals (for
XX CC development and screening of therapeutic agents), also for in vivo or
XX CC ex vivo gene therapy.
XX SQ Sequence 300 AA;

Query Match 100.0%; Score 371; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 5.5e-30;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVQPCRRDSTTCGCPPRHYTFQWNYLCRYCNVLCGEREEARACHATHNRCRCR 60
DB 57 fvgprcrrdspttcgpcprhytqfwnylcrncryncvlgereearachathnrcrcr 116
QY 61 T 61
DB 117 t 117

RESULT 10
W97749
ID W97749 standard; Protein; 300 AA.
XX
XX AC W97749;
XX DT 21-MAY-1999 (first entry)
XX DE Human tumour necrosis factor receptor ZTNFR-5.
XX KW ZTNFR-5; tumour necrosis factor receptor; TNFR; human;
XX OS cell maturation; bone cell regulation.
XX OS Homo sapiens.
XX

```

```

FH Key Location/Qualifiers
FT Peptide 1..23
FT FT /note= "signal peptide"
FT Protein 24..300
FT FT /note= "mature protein"
FT Domain 24..194
FT FT /note= "extracellular domain"
FT Region 49..71
FT FT /note= "cysteine-rich pseudo-repeat 1"
FT Region 72..113
FT FT /note= "cysteine-rich pseudo-repeat 1"
FT Region 114..151
FT FT /note= "cysteine-rich pseudo-repeat 1"
FT Region 152..194
FT FT /note= "cysteine-rich pseudo-repeat 1"
XX WO9904001-A1.
XX PN
XX XX 28-JAN-1999.
XX PD
XX XX 21-JUL-1998; 98WO-US15072.
XX PF
XX XX 21-JUL-1997; 97US-0053203.
XX PR
XX XX (ZYMO ) ZYMOGENETICS INC.
XX PA Farrah TM;
XX PI
XX XX WPI; 1999-132245/11.
XX DR N-PSDB; X07226.
XX DR
XX PT Novel tumour necrosis factor receptor ZTNFR5 - useful for
XX CC regulating maturation of TNF-ligand bearing cells
XX XX Claim 1; Page 84-85; 109pp; English.
XX CC This polypeptide comprises a new, secreted tumour necrosis factor
XX CC receptor (see W97749), designated ZTNFR-5. Novel ZTNFR-5 encoding
XX CC polynucleotides and polypeptides were initially identified by
XX CC querying an expressed sequence tag (EST) database for sequences
XX CC homologous to conserved motifs within the TNF receptor family.
XX CC Based on this search, a contig of 16 ESTs (see X07226) was
XX CC constructed. ZTNFR-5 polypeptides comprise 4 cysteine-rich repeats
XX CC (see also W97750-55) that are homologous to other TNF receptors, in
XX CC particular the soluble, secreted TNF receptor osteoprotegerin.
XX CC ZTNFR-5 polypeptide can be prepared by recombinant methods. The
XX CC polypeptide, especially the extracellular domain, can be used to
XX CC generate a soluble variant of ZTNFR-5. The polypeptides and
XX CC nucleic acids can be used to screen for ligands, agonists and
XX CC antagonists of ZTNFR-5. The polypeptides can be used in bone cell
XX CC regulation and to regulate the maturation of TNF ligand-bearing
XX CC cells such as T- or B-cells, lymphocytes, peripheral blood
XX CC mononuclear cells, polymorphonuclear leukocytes, fibroblasts or
XX CC haematopoietic cells.
XX XX
XX SQ Sequence 300 AA;

```

```

Query Match 100.0%; Score 371; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 5.5e-30;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVQPCRRDSTTCGCPPRHYTFQWNYLCRYCNVLCGEREEARACHATHNRCRCR 60
DB 57 fvgprcrrdspttcgpcprhytqfwnylcrncryncvlgereearachathnrcrcr 116
QY 61 T 61
DB 117 t 117

RESULT 11
W95082

```

ID W95082 standard; Protein; 300 AA.
 AC W95082;
 DT 20-MAY-1999 (first entry)
 XX
 DE Orphan receptor (HUMAN NTR-1) polypeptide.
 XX
 XX HUMAN NTR-1; orphan receptor; osteoprotegerin; OPG; TNFR; human;
 KW tumour necrosis factor receptor; muscle disorder; bone mass; screening;
 KW muscle metabolism; binding agent; cognate ligand.
 XX
 OS Homo sapiens.
 XX
 PN WO9907738-A2.
 XX
 PD 18-FEB-1999.
 XX
 XX 04-AUG-1998; 98WO-US16202.
 PF
 XX 06-AUG-1997; 97US-0054869.
 PR
 XX (PROC) PROCTER & GAMBLE CO.
 PA (REGG-) REGENERON PHARM INC.
 XX
 PI Maslakowski PJ, Morris J, Valenzuela DM;
 XX
 DR WPI: 1999-167365/14.
 DR N-PSDB; X22300.
 XX
 PT Novel orphan human receptor polypeptide and nucleic acid - useful as
 PT diagnostic reagents and for treatment of muscle disorders
 PS
 XX Claim 7; Page 21; 23pp; English.
 CC This represents a HUMAN NTR-1 polypeptide, a novel orphan receptor. The
 CC protein is related to osteoprotegerin (OPG) and to tumour necrosis factor
 CC receptor (TNFR). Host cells transfected with a vector comprising the
 CC HUMAN NTR-1 nucleic acid are used for the recombinant expression of the
 CC protein. HUMAN NTR-1 proteins and antibodies immuno specific for the
 CC protein are useful for diagnosis and treatment of humans and animals,
 CC especially muscle disorders, as the receptor is involved in regulation of
 CC bone mass and muscle metabolism. HUMAN NTR-1 receptors are also useful
 CC for screening for novel binding agents, and cognate ligands, which may be
 CC used to treat disorders associated with HUMAN NTR-1 imbalance.
 XX
 SQ Sequence 300 AA;
 Query Match 100.0%; Score 371; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 5.5e-30;
 Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FVQPCRRDSTTCGCPPRHYTFQWNYLCRCRYCNVLCGEREEARACHATHNRCRCR 60
 DB 57 fvgprcrrdspttcgpcpprhhytqfwnylcrfrcrycnvlgereearachathnrcrcr 116
 QY 61 T 61
 DB 117 t 117
 RESULT 12
 Y42185
 ID Y42185 standard; Protein; 273 AA.
 AC Y42185;
 XX
 XX 17-DEC-1999 (first entry)
 DT
 XX Human mFLINT #2 protein sequence.
 DE
 XX Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;

KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
 KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
 KW reperfusion-associated injury; aplastic anaemia; differentiation;
 KW growth; myelodysplastic syndrome; pancytopenic condition;
 KW myocardial ischaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO99504113-A2.
 XX
 PD 07-OCT-1999.
 XX
 PF 30-MAR-1999; 99WO-US06797.
 PR
 XX 30-MAR-1998; 98US-0079856.
 PR 20-MAY-1998; 98US-0086074.
 PR 09-SEP-1998; 98US-0099643.
 PR 17-DEC-1998; 98US-0112577.
 PR 18-DEC-1998; 98US-0112703.
 PR 18-DEC-1998; 98US-0112933.
 PR 22-DEC-1998; 98US-0113407.
 XX
 PA (ELIL) LILLY & CO ELI.
 XX
 PI Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
 PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
 PI Song HY, Wang J, Wu X, Zuckerman SH;
 XX
 DR WPI: 1999-591319/50.
 DR N-PSDB; Z25378.
 XX
 PT Use of mature FLINT for treating acute liver failure, inflammation,
 PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
 PT and proinflammatory activity
 XX
 PS Example 2; Fig 4; 99pp; English.
 CC The present invention describes therapeutic applications of mature FLINT
 CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
 CC (mFLINT), which is a member of the tumour necrosis factor receptor
 CC superfamily, is used for treating acute liver failure, inflammation of
 CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
 CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
 CC injury or disorder such as hypercoagulation (including use with
 CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
 CC or disorder, Type I diabetes, cancer, cell damage or damage to an
 CC innocent bystander tissue that is induced by a chemotherapeutic agent or
 CC therapeutic irradiation, treating haematopoietic progenitor cells that
 CC have been exposed to therapeutic radiation or chemotherapy, aplastic
 CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
 CC also used for promoting the growth or differentiation of a haematopoietic
 CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
 CC resulting from abnormal myocardial ischaemia. The present sequence
 CC represents human mFLINT.
 XX
 SQ Sequence 273 AA;
 Query Match 98.7%; Score 366; DB 20; Length 273;
 Best Local Similarity 100.0%; Pred. No. 1.6e-29;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 FVQPCRRDSTTCGCPPRHYTFQWNYLCRCRYCNVLCGEREEARACHATHNRCRCR 60
 DB 28 fvgprcrrdspttcgpcpprhhytqfwnylcrfrcrycnvlgereearachathnrcrcr 87
 RESULT 13
 Y42183
 ID Y42183 standard; Protein; 302 AA.
 XX
 XX Y42183;
 XX

DT 17-DEC-1999 (first entry)
DE Human FLINT #2 protein sequence.
XX
XX Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
KW reperfusion-associated injury; aplastic anaemia; differentiation;
KW growth; myelodysplastic syndrome; pancytopenic condition;
KW myocardial ischaemia.
XX
XX Homo sapiens.
OS
XX W099504113-A2.
XX
XX 07-OCT-1999.
XX
XX 30-MAR-1999; 99WO-US06797.
XX
XX 30-MAR-1998; 98US-0079856.
PR 20-MAY-1998; 98US-0086074.
PR 09-SEP-1998; 98US-0099643.
PR 17-DEC-1998; 98US-0112577.
PR 18-DEC-1998; 98US-0112703.
PR 18-DEC-1998; 98US-0112933.
PR 22-DEC-1998; 98US-0113407.
XX
XX (ELIL) LILLY & CO ELI.
XX
XX Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
PI Hui XY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
PI Song HY, Wang X, Wu X, Zuckerman SH;
XX
XX WPI; 1999-591319/50.
DR N-PSDB; 225376.
XX
XX Use of mature FLINT for treating acute liver failure, inflammation,
PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
PT and proinflammatory activity
XX
XX Example 2; Fig 2; 99pp; English.
XX
XX The present invention describes therapeutic applications of mature FLINT
CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
CC (mFLINT), which is a member of the tumour necrosis factor receptor
CC superfamily, is used for treating acute liver failure, inflammation of
CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
CC injury or disorder such as hypercoagulation (including use with
CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
CC or disorder, type I diabetes, cancer, cell damage or damage to an
CC innocent bystander tissue that is induced by a chemotherapeutic agent or
CC therapeutic irradiation, treating haematopoietic progenitor cells that
CC have been exposed to therapeutic radiation or chemotherapy, aplastic
CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
CC also used for promoting the growth or differentiation of a haematopoietic
CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
CC resulting from abnormal myocardial ischaemia. The present sequence
XX represents human FLINT.
XX
XX Sequence 302 AA;
SQ

Query Match 98.7%; Score 366; DB 20; Length 302;
Best Local Similarity 100.0%; Pred. NO. 1.7e-29;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVQPCRRDPTTCGPPPHRYTQFNWYLCRCRYCNVLCGEREEARACHATHNRCRCR 60
|||||
Db 57 fvqpcrrdspttcgpcpprhycqfwnylcrvcnvcgceeeearachathnrcr 116
|||||

RESULT 14

Y77458
ID Y77458 standard; Protein; 300 AA.
XX
XX Y77458;
AC
XX 05-JUN-2000 (first entry)
DT
XX Human TNF receptor-like protein, HDTEA84.
DE
XX
XX TNF receptor family; tumour necrosis factor; HDTEA84; HSLJD37R;
KW Rank-like protein; RANKL; immune disorder; inflammation; allergy;
KW immunosuppressant; antiarthritic; antirheumatoid; antiinflammatory;
KW dermatological; antithyroid.
XX
XX Homo sapiens.
OS
XX W0200001817-A2.
XX
XX 13-JAN-2000.
XX
XX 06-JUL-1999; 99WO-US12366.
XX
XX 06-JUL-1998; 98US-0110938;
PR 13-JUL-1998; 98US-0114466;
PR 23-JUL-1998; 98US-0093897;
PR 12-AUG-1998; 98US-0132968;
PR 18-AUG-1998; 98US-0136214;
PR 11-SEP-1998; 98US-0099999;
XX
XX (SCHE) SCHERING CORP.
XX
XX Bates ERM, Lebecque SJE, Murphy EE, Mattson JD, Gorman DM;
PI Hedrick JA, Wang L, Zlotnik A, Murgolo NJ, Greene JR, Johnston JA;
PI Bazan JF, Mahony D, Lees EM;
XX
XX WPI; 2000-171015/15.
DR N-PSDB; 292404.
XX
XX New isolated mammalian genes, used to develop products for treating
PT e.g. immune, inflammatory or allergic abnormalities, cancers or
PT degenerative conditions
XX
XX Claim 24; Page 157; 218pp; English.
XX
XX The invention relates to a number of primate and/or rodent proteins, and
CC the genes which encode them. The invention encompasses human dendritic
CC cell prostaglandin transporter (DC-PGT); the TNF (tumour necrosis
CC factor) receptor family-related proteins HDTEA84, HSLJD37R and RANKL;
CC human CC chemokine HCC5; human deubiquitinating proteins Dub11 and Dub
CC 12; human MD-1 and human and murine MD-2 proteins, which exhibit the
CC properties of ligands for proteins comprising a leucine-rich motif (LRR);
CC human cyclin E2; cDNAs encoding these proteins; and antibodies against
CC these proteins. The proteins can be used for modulating the physiology or
CC development of a cell. They can be used to mediate uptake of substrates
CC (e.g., prostaglandin-like molecules), to modulate or mediate cellular
CC interactions (e.g., induce or prevent trafficking, proliferation, or
CC differentiation of cells), or are intracellular proteins which are
CC important in various cellular processes such as the deubiquitination of
CC proteins or cell cycle regulation. The products can be used for
CC treating medical conditions such as immune, inflammatory or allergic
CC disorders, or abnormal cellular proliferation, for example, cancers or
CC degenerative conditions. They can be used to modulate immune responses in
CC disease states e.g., autoimmune disorders, including rheumatoid
CC arthritis, systemic lupus erythematosus, Hashimoto's autoimmune
CC thyroiditis, as well as acute and chronic inflammatory responses in which
CC T cell activation, expansion, and/or immunological T cell memory play an
CC important role. Sequences Y77458-Y77461 and Y77465-Y77468 represent
CC TNF receptor family-related proteins. Y77458 is the human protein
CC HDTEA84, Y77459-Y77461 are human HSLJD37R proteins, Y77465 is
CC murine Rank-like protein RANKL, and Y77466-Y77468 are human RANKL
CC proteins.
XX
XX Sequence 300 AA;
SQ

Query Match 96.2%; Score 357; DB 21; Length 300;
Best Local Similarity 96.7%; Pred. No. 1.4e-28;
Matches 59; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 FVQPCRRDSPITGCPGPPRHVTQFWNYLERCRYCNVLCGEREEARACHATHNRCRCR 60
Db 57 fvqpcrrdspmtgpcpprhvtqfwnylercrycnvlgereearachathnrcrcr 116

QY 61 T 61

Db 117 t 117

RESULT 15

Y22222
ID Y22222 standard; Protein; 153 AA.

XX AC Y22222;

XX 16-SEP-1999 (first entry)

XX Human TNFR superfamily soluble receptor protein sequence.

XX TNFL1; human; TNFR superfamily; tumour necrosis factor ligand; TNF;
KW tumour necrosis factor receptor; TNFR superfamily; cell proliferation;
KW cell differentiation; cytokine production; immunoglobulin; hyperplasia;
KW apoptosis inducer; activated T cell; autoimmune disease; inhibitor;
KW myasthenia gravis; insulin-dependent diabetes mellitus; endotoxin shock;
KW rheumatoid arthritis; multiple sclerosis; systemic lupus erythematosus;
KW tumour; proliferative disorder; neoplasia; dysplasia; immunocompetence;
KW lymphoid organogenesis; bacterial resistance; contact hypersensitivity;
KW delayed type sensitivity; therapy.

XX OS Homo sapiens.

XX PN W09933980-A2.

XX PD 08-JUL-1999.

XX PF 22-DEC-1998; 98WO-US27474.

XX PR 16-DEC-1998; 98US-0212270.

XX PR 30-DEC-1997; 97US-0068959.

XX PA (CHIR) CHIRON CORP.

XX PI Kassam A, Lamson G, Pot D, Tribouley C;

XX DR WPI; 1999-405508/34.

XX DR N-PSDB; X84621.

XX New tumour necrosis factor ligands, useful for induction of cell
PT death and/or proliferation of cells

XX Claim 1; Page 61; 69pp; English.

XX This sequence represents a tumour necrosis factor receptor (TNFR)
CC superfamily soluble protein of the invention. The invention also relates
CC to tumour necrosis factor (TNF) ligand (TNFL) family proteins. The TNFL
CC proteins play regulatory roles in cell proliferation and/or
CC differentiation, e.g. they can induce production of cytokines,
CC immunoglobulins, etc. A variety of diseases can be treated by modulating
CC the activity of TNFL proteins, e.g. they can induce apoptosis of
CC activated T cells but rescue resting T cell from apoptosis. TNFL
CC polypeptides can therefore be used to treat autoimmune diseases, such as
CC myasthenia gravis, insulin-dependent diabetes mellitus, rheumatoid
CC arthritis, multiple sclerosis, and systemic lupus erythematosus. TNFL
CC proteins also have tumour stimulating properties, so tumours can be
CC treated by inhibiting the expression or activity of TNFL. Other
CC proliferative disorders, such as neoplasias, dysplasias, and hyperplasia
CC can also be treated using TNFL inhibitors. The TNFL polypeptides and

CC polynucleotides can also be used to enhance or decrease TNF activity,
CC thus providing therapeutic benefits such as induction of cell death,
CC lymphoid organogenesis, or host bacterial resistance, and inhibition of
CC endotoxin shock, contact hypersensitivity, delayed type sensitivity or
CC immunocompetence of a transplant recipient. TNF and its receptors play a
CC major role in host defence and immunosurveillance. As such, there is a
CC need to identify new members of TNFR families. This invention provides
CC this need.

XX Sequence 153 AA;

Query Match

Best Local Similarity 52.0%; Score 193; DB 20; Length 153;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 LERCRCNVLCGEREEARACHATHNRCRCRT 61

Db 1 lercrcnvlgereearachathnrcrcrt 33

Search completed: January 30, 2001, 16:45:22
Job time: 582 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:50:19 ; Search time 149.64 seconds
(without alignments)
27.679 Million cell updates/sec

Title: US-09-518-931-2_COPY_57_117

Perfect score: 371

Sequence: 1 FVQPCRDSPPTGCPGPPR.....REERACHATHNRACRRT 61

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_66: *
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	129	34.8	459	2	I48854	gene murine tumour
2	128	34.5	461	1	A35356	tumor necrosis fac
3	126	34.0	348	2	T28623	hypothetical prote
4	126	34.0	349	2	D72175	G2R protein - vari
5	126	34.0	349	2	D36858	gene G4R protein -
6	126	34.0	474	2	B38634	tumor necrosis fac
7	106	28.6	325	2	B43692	T2 protein - rabbi
8	103.5	27.9	435	2	I34182	tumor necrosis fac
9	97	26.1	326	1	GVQZML	T2 protein - myxom
10	96	25.9	277	2	I37552	OX40 homolog - hum
11	95.5	25.7	416	1	JN0006	nerve growth facto
12	88	23.7	271	2	S12783	OX40 antigen precu
13	86.5	23.3	324	2	JC2395	Fas antigen precu
14	85.5	23.0	327	2	A46484	apoptosis-mediati
15	83	22.4	272	2	I48700	gene ox40 protein
16	79.5	21.4	314	2	I37383	FAS soluble protei
17	79.5	21.4	335	2	A40036	apoptosis-mediati
18	78.5	21.2	256	2	B32393	T-cell antigen 4-1
19	78	21.0	260	1	A46517	CD27 antigen precu
20	74	19.9	255	2	I38426	4-1BB - human
21	74	19.9	255	2	J70752	lymphocyte activat
22	74	19.9	4660	2	T42737	gp330 protein prec
23	72.5	19.5	425	1	A26431	nerve growth facto
24	72.5	19.5	427	1	GOHUN	nerve growth facto
25	71	19.1	461	2	JC4302	tumor necrosis fac
26	70.5	19.0	1299	2	T43251	furin (EC 3.4.21.7
27	69.5	18.7	454	1	GMST1	tumor necrosis fac
28	69.5	18.7	454	2	I57826	tumor necrosis fac
29	68.5	18.5	225	2	A31926	transcription acti

ALIGNMENTS

RESULT 1

I48854

gene murine tumour necrosis factor receptor 2 protein - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 23-Jul-1999
C:Accession: I48854
R:Powell, E.E.; Wicker, L.S.; Peterson, L.B.; Todd, J.A.
Mamm. Genome 5, 726-727, 1994
A:Title: Allelic variation of the type 2 tumor necrosis factor receptor gene.
A:Reference number: I48854; MUID:95178848
A:Accession: I48854

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-459 <RES>
A:Cross-references: EMBL:X76401; NID:g433830; PIDN:CAA53981.1; PID:g433831
C:Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology
F:151-188/Domain: NGF receptor repeat homology <NGF>

Query Match 34.8%; Score 129; DB 2; Length 459;
Best Local Similarity 37.3%; Pred. No. 6e-06;
Matches 22; Conservative 5; Mismatches 32; Indels 0; Gaps 0;

Qy 1 FVQPCRDSPPTGCPGPPRHYTFQWNYLCRGEEERACHATHNRACRC 59
Db 48 YVKHFCNKTSDTVCADCEASMYTQVNOFRCLSCSSCSSTQDVETRACTKOONRVAC 106

RESULT 2

A35356

tumor necrosis factor receptor type 2 precursor - human
N:Alternate names: 75K tumor necrosis factor receptor
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A35356; A36475; A48416; A36007; A23666; B35010; I38094
R:Smith, C.A.; Davis, T.; Anderson, D.; Solam, L.; Beckmann, M.P.; Jerzy, R.; Dower, Science 248, 1019-1023, 1990
A:Title: A receptor for tumor necrosis factor defines an unusual family of cellular a
A:Reference number: A35356; MUID:90260639
A:Accession: A35356
A:Status: preliminary
A:Molecule type: mRNA

A:Residues: 1-461 <SM1>

A:Cross-references: GB:M32315; NID:g189185; PIDN:AA559929.1; PID:g189186
R:Kohn, T.; Brewer, M.T.; Baker, S.L.; Schwartz, P.E.; King, M.W.; Hale, K.K.; Squir Proc. Natl. Acad. Sci. U.S.A. 87, 8331-8335, 1990
A:Title: A second tumor necrosis factor receptor gene product can shed a naturally oc
A:Reference number: A36475; MUID:91045991
A:Accession: A36475
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-195,'R',197-461 <KOH>

db 67 CT K T S D T V C D S C E F N S T Y T O I W N J V D E C I S C C C S D C S C D O V E T O A C M E R F O N S T C M E

A;Molecule type: DNA
A;Residues: 1-349 <BLI>

Query Match	23.78;	Score 88;	DB 2;	Length 271;
Best Local Similarity	33.3%;	Pred. No. 0.053;		
Matches 19;	Conservative	6;	Mismatches 28;	Indels 4;
Gaps				
Qy	6	CRDSPTTCGCPRHYYTFQWNY	--LERCRCYNLCGEREEEARACHATHNRACRCR	60
Db	51	CDHWRTDYCHCEPGFYNEAYNYDTCKCTOCNHRSG	--SELKONCTPTEDTVQCR	105

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model
Run on: January 30, 2001, 17:05:55 ; Search time 80.04 Seconds
(without alignments)
24.612 Million cell updates/sec

Title: US-09-518-931-2_COPY_57_117
Perfect score: 371
Sequence: 1 FVQPCRRDPTTCGCPPR.....REEEARACHATHNACRCRT 61

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	128	34.5	461	1 TNR2_HUMAN	P20333 homo sapien
2	126	34.0	349	1 VC22_VARV	P34015 variola vir
3	126	34.0	474	1 TNR2_MOUSE	P25119 mus musculus
4	112.5	30.3	415	1 TNR2_MOUSE	P50284 mus musculus
5	106	28.6	325	1 VT2_SFVKA	P25943 shope fibro
6	103.5	27.9	435	1 TNR2_HUMAN	P36941 homo sapien
7	97	26.1	326	1 VT2_MXVVL	P29825 myxoma viru
8	95.5	25.9	277	1 OX40_HUMAN	P43489 homo sapien
9	88	23.7	271	1 NGFR_CHICK	P18519 gallus gall
10	88	23.7	271	1 OX40_RAT	P15725 rattus norv
11	86.5	23.3	324	1 FASA_PIG	O77736 sus scrofa
12	86.5	23.3	324	1 FASA_RAT	Q63199 rattus norv
13	85.5	23.0	327	1 FASA_MOUSE	P25446 mus musculus
14	83	22.4	272	1 OX40_MOUSE	P47741 mus musculus
15	80	21.6	471	1 TNR1_BOVIN	O19131 bos taurus
16	79.5	21.4	335	1 FASA_HUMAN	P25445 homo sapien
17	79	21.3	323	1 FASA_BOVIN	P51867 bos taurus
18	78.5	21.2	256	1 41BB_MOUSE	P20334 mus musculus
19	74	19.9	255	1 41BB_HUMAN	Q07011 homo sapien
20	74	19.9	260	1 CD27_HUMAN	P26842 homo sapien
21	74	19.9	4660	1 LRP2_RAT	P98158 rattus norv
22	72.5	19.5	425	1 NGFR_RAT	P07174 rattus norv
23	72.5	19.5	427	1 NGFR_HUMAN	P08138 homo sapien
24	71	19.1	461	1 TNR1_PIG	P50555 sus scrofa
25	69.5	18.7	454	1 TNR1_MOUSE	P25118 mus musculus
26	68.5	18.5	225	1 ACE1_YEAST	P53315 saccharomyc
27	68	18.3	250	1 CD27_MOUSE	P41272 mus musculus
28	68	18.3	455	1 TNR1_HUMAN	P19438 homo sapien
29	65	17.5	461	1 TNR1_RAT	P22934 rattus norv
30	65	17.5	677	1 SPB7_DICDI	P54643 dictyosteli
31	64	17.3	164	1 VEGF_CAVPO	P26617 cavia porce
32	64	17.3	612	1 UNC6_CAEEL	P34710 caenorhabdi
33	62.5	16.8	704	1 KPCL_CAEEL	P34722 caenorhabdi

34	62.5	16.8	1173	1	TSPL_XENLA	P35448 xenopus lae
35	62.5	16.8	4655	1	LRP2_HUMAN	P98164 homo sapien
36	62	16.7	1700	1	BAR3_CHITE	Q03376 chironomus
37	61.5	16.6	122	1	VC22_VACCC	P21106 vaccinia vi
38	61	16.4	190	1	VEGF_BOVIN	P15691 bos taurus
39	61	16.4	190	1	VEGF_PIG	P49151 sus scrofa
40	61	16.4	595	1	CD30_HUMAN	P28908 homo sapien
41	60.5	16.3	1745	1	YL51_CAEEL	P34431 caenorhabdi
42	60	16.2	1104	1	NFX1_HUMAN	Q12986 homo sapien
43	59.5	16.0	269	1	CD40_BOVIN	Q28203 bos taurus
44	59.5	16.0	289	1	CD40_MOUSE	P27512 mus musculu
45	59.5	16.0	909	1	CT1A_FUSO	P52958 fusarium so

ALIGNMENTS

RESULT	1
TNR2_HUMAN	
ID	TNR2_HUMAN STANDARD; PRT; 461 AA.
AC	P20333;
DT	01-FEB-1991 (Rel. 17, Created)
DT	01-AUG-1991 (Rel. 19, Last sequence update)
DT	01-OCT-2000 (Rel. 40, Last annotation update)
DE	TUMOR NECROSIS FACTOR RECEPTOR 2 PRECURSOR (TUMOR NECROSIS FACTOR BINDING PROTEIN 2) (TBPLI) (P80) (TNF-R2) (P75) (CD120B) (ETANERCEPT).
GN	TNFRSF1B OR TNFR2 OR TNFR.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=90260639; PubMed=2160731;
RA	Smith C.A., Davis T., Anderson D., Solam L., Beckmann M.P., Jerzy R., Dower S.K., Cosman D., Goodwin R.G.;
RT	"A receptor for tumor necrosis factor defines an unusual family of cellular and viral proteins.";
RL	Science 248:1019-1023(1990).
RN	[2]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=91045991; PubMed=2172983;
RA	Kohn T., Brewer M.T., Baker S.L., Schwartz P.E., King M.W., Hale K.K., Squires C.H., Thompson R.C., Vannice J.L.;
RT	"A second tumor necrosis factor receptor gene product can shed a naturally occurring tumor necrosis factor inhibitor.";
RL	Proc. Natl. Acad. Sci. U.S.A. 87:8331-8335(1990).
RN	[3]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=96299745; PubMed=86611109;
RA	Beltlinger C.P., White P.S., Maris J.M., Sulman E.P., Jensen S.J., Lepaslier D., Stallard B.J., Goeddel D.V., Desauvage F.J., Brodeur G.M.;
RT	"Physical mapping and genomic structure of the human TNFR2 gene.";
RL	Genomics 35:94-100(1996).
RN	[4]
RP	SEQUENCE OF 116-461 FROM N.A. AND PARTIAL SEQUENCE.
RX	MEDLINE=90349572; PubMed=2166946;
RA	Heller R.A., Song K., Onasch M.A., Fischer W.H., Chang D., Ringold G.M.;
RT	"Complementary DNA cloning of a receptor for tumor necrosis factor and demonstration of a shed form of the receptor.";
RL	Proc. Natl. Acad. Sci. U.S.A. 87:6151-6155(1990).
RN	[5]
RP	SEQUENCE OF 27-31.
RX	MEDLINE=90110215; PubMed=2153136;
RA	Engelmann H., Novick D., Wallach D.;
RT	"Two tumor necrosis factor-binding proteins purified from human urine. Evidence for immunological cross-reactivity with cell surface tumor necrosis factor receptors.";
RL	J. Biol. Chem. 265:1531-1536(1990).
RN	[6]
RP	SEQUENCE OF 22-40; 65-69; 136-141; 300-306 AND 346-362.
RX	MEDLINE=91056048; PubMed=2173696;

RA Loetscher H., Schlaeager E.J., Lahm H.-W., Pan Y.-C.E., Lesslauer W.,
 RA Brockhaus M.;
 RT "Purification and partial amino acid sequence analysis of two
 RT distinct tumor necrosis factor receptors from HL60 cells.";
 RT J. Biol. Chem. 265:20131-20138(1990).
 RN [7]
 RP CHARACTERIZATION.
 RX MEDLINE-93016040; PubMed-1328224;
 RA Pennica D., Lam V.T., Mize N.K., Weber R.F., Lewis M., Fendly B.M.,
 RA Lipari M.T., Goeddel D.V.;
 RT "Biochemical properties of the 75-kDa tumor necrosis factor receptor.
 RT Characterization of ligand binding, internalization, and receptor
 RT phosphorylation.";
 RT J. Biol. Chem. 267:21172-21178(1992).
 RN [8]
 RX X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 419-428 IN TRAF2 COMPLEX.
 RX MEDLINE-99221490; PubMed-10206649;
 RA Park Y.C., Burkitt V., Villa A.R., Tong L., Wu H.;
 RT "Structural basis for self-association and receptor recognition of
 RT human TRAF2.";
 RL Nature 398:533-538(1999).
 CC -!- FUNCTION: RECEPTOR FOR TNF-ALPHA. HIGH AFFINITY FOR TNA-ALPHA AND
 CC APPROXIMATELY 5-FOLD LOWER AFFINITY FOR TNF-BETA.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -!- PTM: PHOSPHORYLATED; MAINLY ON SERINE RESIDUES WITH A VERY LOW
 CC LEVEL ON THREONINE RESIDUES.
 CC -!- PHARMACEUTICAL: AVAILABLE UNDER THE NAME ENBREL (IMMUNEX AND
 CC WYETH-AVERT). USED TO TREAT MODERATE TO SEVERE RHEUMATOID
 CC ARTHRITIS (RA). ENBREL CONSIST OF THE EXTRACELLULAR LIGAND-BINDING
 CC PORTION OF TNFR2 LINKED TO AN IMMUGLOBULIN FC CHAIN. IT BINDS TO
 CC TNF-ALPHA AND BLOCKS ITS INTERACTIONS WITH RECEPTORS.
 CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 CC -!- DATABASE: NAMB-PHOW; NOTE=CD guide CD120b entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd120b.htm".
 CC -!- DATABASE: NAMB-Enbrel; NOTE=Clinical information on Enbrel;
 CC WWW="http://www.enbrelinfo.com/".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M32315; AAC59929.1; -;
 CC EMBL; M35857; AAC63262.1; -;
 CC EMBL; U52165; AAC50622.1; -;
 CC EMBL; U52156; AAC50622.1; JOINED.
 CC EMBL; U52157; AAC50622.1; JOINED.
 CC EMBL; U52158; AAC50622.1; JOINED.
 CC EMBL; U52159; AAC50622.1; JOINED.
 CC EMBL; U52160; AAC50622.1; JOINED.
 CC EMBL; U52161; AAC50622.1; JOINED.
 CC EMBL; U52162; AAC50622.1; JOINED.
 CC EMBL; U52163; AAC50622.1; JOINED.
 CC EMBL; U52164; AAC50622.1; JOINED.
 CC EMBL; M55994; AAA36755.1; -;
 CC PIR; A35356; A35356.
 CC PIR; A36007; A36007.
 CC PIR; A36475; A36475.
 CC PIR; B35010; B35010.
 CC PIR; A23666; A23666.
 CC PDB; 1CA9; 12-APR-99.
 CC MIM; 191191; -;
 CC INTERPRO: IPR001368; -;
 CC PFAM: PF00020; TNFR_C6; 4.
 CC PROSITE: PS00652; TNFR_NGFR_1; 2.
 CC PROSITE: PS00050; TNFR_NGFR_2; 4.
 KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal;
 KW Phosphorylation; Pharmaceutical; 3D-structure.
 FT SIGNAL 1 22
 FT CHAIN 23 461 TUMOR NECROSIS FACTOR RECEPTOR 2.

FT DOMAIN 23 257 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 258 287 POTENTIAL.
 FT DOMAIN 288 461 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 39 201 4 X TNFR-CYS.
 FT REPEAT 39 76 TNFR-CYS 1.
 FT REPEAT 77 118 TNFR-CYS 2.
 FT REPEAT 119 162 TNFR-CYS 3.
 FT REPEAT 163 201 TNFR-CYS 4.
 FT DISULFID 40 53 BY SIMILARITY.
 FT DISULFID 54 67 BY SIMILARITY.
 FT DISULFID 57 75 BY SIMILARITY.
 FT DISULFID 78 93 BY SIMILARITY.
 FT DISULFID 96 110 BY SIMILARITY.
 FT DISULFID 100 118 BY SIMILARITY.
 FT DISULFID 120 126 BY SIMILARITY.
 FT DISULFID 134 143 BY SIMILARITY.
 FT DISULFID 137 161 BY SIMILARITY.
 FT CARBOHYD 171 171 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 193 193 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 141 141 R -> P (IN REF. 4).
 FT CONFLICT 196 196 R -> M (IN REF. 1 AND 3).
 FT CONFLICT 363 363 A -> T (IN REF. 4).
 SQ SEQUENCE 461 AA; 48316 MW; 603B580ECD67636F CRC64;
 Query Match 34.5%; Score 128; DB 1; Length 461;
 Best Local Similarity 38.2%; Pred. No. 4.9e-07;
 Matches 21; Conservative 5; Mismatches 29; Indels 0; Gaps 0;
 OY 6 CRDSDTTTCGPPRHYYTFWNYLRCRCVNLGGEREEAEARACHATHNRCRCR 60
 DB 67 CRTKSTVDCSDSTYTQLMWNVPECLSCGRSCSDQVETQACTREQNRICR 121
 RESULT 2
 VC22_VARV STANDARD; PRT; 349 AA.
 AC VC22_VARV
 AC P34015;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE PROTEIN C22/B28 HOMOLOG.
 GN G4R.
 OS Variola virus.
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Orthopoxvirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=INDIA-1967 / ISOLATE IND3;
 RX MEDLINE-93202281; PubMed-8384129;
 RA Shchelkunov S.N., Blinov V.M., Sandakhchiev L.S.;
 RT "Genes of variola and vaccinia viruses necessary to overcome the host
 RT protective mechanisms."
 RL FEBS Lett. 319:80-83(1993).
 CC -!- SIMILARITY: CONTAINS 2 LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGIONS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X69198; CAA49137.1; -;
 CC EMBL; X67117; CAA47540.1; -;
 CC PIR; D36858; D36858.
 CC PIR; S35987; S35987.
 CC PIR; S46888; S46888.
 CC HSSP; PI9438; INCF.
 CC INTERPRO: IPR001368; -;
 CC PFAM; PF00020; TNFR_C6; 2.


```

DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
DR Repeat.
FT DOMAIN 31 108 2 X TNFR-CYS.
FT REPEAT 31 66 TNFR-CYS 1.
FT REPEAT 67 108 TNFR-CYS 2.
FT REPEAT 67 108 TNFR-CYS 2.
SQ SEQUENCE 349 AA; 39189 MW; D45D40B5C6E780EF CRC64;

Query Match 34.0%; Score 126; DB 1; Length 349;
Best Local Similarity 37.3%; Pred. No. 6.2e-07;
Matches 22; Conservative 7; Mismatches 30; Indels 0; Gaps 0;

QY 1 FVQRCRDSPPTGCPGPPRHYYTFQWYLEKRCYCNVLCGEREBEERACHATHNRACRC 59
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 52 YASRLCDSKNTQCTPCGSGTFTSRNNHLPACLSCNGRCNSQVETRSCNTTHNRICEC 110

RESULT 3
TNR2_MOUSE STANDARD; PRT; 474 AA.
ID TNR2_MOUSE AC P25119; P97893;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR 2 PRECURSOR (TNF-R2) (P75).
OS TNFRSF1B OR TNFR2 OR TNFR-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-91187885; PubMed=1849278;
RA Lewis M., Tartaglia L.A., Lee A., Bennett G.L., Rice G.C.,
RA Wong G.H., Chen E.Y., Goeddel D.V.;
RT "Cloning and expression of cDNAs for two distinct murine tumor
RT necrosis factor receptors demonstrate one receptor is species
RT specific.";
RT Proc. Natl. Acad. Sci. U.S.A. 88:2830-2834(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-91246168; PubMed=1645445;
RA Goodwin R.G., Anderson D., Jerzy R., Daviss T., Brannan C.I.,
RA Copeland N.G., Jenkins N.A., Smith C.A.;
RT "Molecular cloning and expression of the type 1 and type 2 murine
RT receptors for tumor necrosis factor.";
RL Mol. Cell. Biol. 11:3020-3026(1991).
RN [3]
RP SEQUENCE OF 1-26 FROM N.A.
RC STRAIN=NOD;
RA Jacob C.O., Liu J.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 1-22 FROM N.A.
RC TISSUE=Liver;
RA Kissonerghis M., Fellows R., Feldmann M., Chernaiovsky Y.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: RECEPTOR FOR TNF-ALPHA.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M60469; AAA39752.1; -
DR EMBL; M59378; AAA0463.1; -
DR EMBL; U39488; AAA85021.1; -
DR EMBL; X87128; CAA60618.1; -

```



```
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=93252381; PubMed=8486360;
RA Baens M., Chaffanet M., Cassiman J.J., den Berghe H., Marynen P.;
RT "Construction and evaluation of a hncDNA library of human l2p
RL transcribed sequences derived from a somatic cell hybrid.";
RN Genomics 16:214-218(1993).
RN [2]
RP FUNCTION.
RX MEDLINE=94225209; PubMed=8171323;
RA Crowe P.D., van Arsdale T.L., Walter B.N., Ware C.F., Hession C.,
RA Ehrenfelds B., Browning J.L., Din W.S., Goodwin R.G., Smith C.A.;
RT "A lymphotoxin-beta-specific receptor.";
RL Science 264:707-710(1994).
CC -!- FUNCTION: RECEPTOR FOR THE LYMPHOTOXIN-BETA. POSSIBLE FUNCTION IN
CC IMMUNE DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L04270; AAA36757.1; -.
DR HSSP; P25942; 1CDF.
DR MIM; 600979; -.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00050; TNFR_NGFR_2; 3.
KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 30
FT CHAIN 31 435
FT DOMAIN 31 227
FT TRANSFAM 228 247
FT DOMAIN 249 435
FT DOMAIN 42 211
FT REPEAT 42 81
FT REPEAT 82 124
FT REPEAT 125 168
FT REPEAT 169 211
FT DISULFID 43 58
FT DISULFID 59 72
FT DISULFID 62 80
FT DISULFID 83 98
FT DISULFID 101 116
FT DISULFID 104 124
FT DISULFID 126 132
FT DISULFID 139 148
FT DISULFID 142 167
FT DISULFID 170 185
FT CARBOHYD 40 40
FT CARBOHYD 177 177
FT CARBOHYD 177 177
SQ SEQUENCE 435 AA; 46709 MW; 624626E6022F656F CRC64;

Query Match 27.9%; Score 103.5; DB 1; Length 435;
Best Local Similarity 33.3%; Pred. No. 0.00021;
Matches 21; Conservative 7; Mismatches 30; Indels 5; Gaps 2;

Qy 1 FVQPCRDSPPTCGPCPPRHVTFWNYL---ERCYCNVLCGEREEERACHATHNRAC 57
:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 67 YVSAKSRIRDTVCATCAENSYNHNWYLTICQLCRPCDPVNG--LEEIAPTCKSRKQTQ 124
:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Qy 58 RCR 60
||:
```

```
Db 125 RCQ 127

RESULT 7
VT2_MYXVL STANDARD; PRT; 326 AA.
AC P29825;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 36, Last annotation update)
DE TUMOR NECROSIS FACTOR SOLUBLE RECEPTOR PRECURSOR (PROTEIN T2).
GN T2.
OS Myxoma virus (strain Lausanne).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Leporipoxvirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91335768; PubMed=1651597;
RA Upton C., Macen J.L., Schreiber M., McFadden G.;
RT "Myxoma virus expresses a secreted protein with homology to the tumor
RT necrosis factor receptor gene family that contributes to viral
RT virulence.";
RL Virology 184:370-382(1991).
CC -!- FUNCTION: BINDS TO TNF-ALPHA AND BETA. PROBABLY PREVENTS TNF TO
CC REACH CELLULAR TARGET AND THEREBY DEAMPENING THE POTENTIAL
CC ANTIVIRAL EFFECTS OF THE CYTOKINE.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M95181; AAA46632.1; -.
DR EMBL; A23729; CAA01688.1; -.
DR PIR; A40566; GQVZML.
DR HSSP; P19438; 1TNR.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 2.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00050; TNFR_NGFR_2; 2.
KW Receptor; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 16
FT CHAIN 17 326
FT DOMAIN 27 186
FT REPEAT 27 62
FT REPEAT 63 104
FT REPEAT 105 147
FT REPEAT 148 186
FT REPEAT 186 205
FT CARBOHYD 66 66
FT CARBOHYD 181 181
FT CARBOHYD 205 205
FT CARBOHYD 238 238
SQ SEQUENCE 326 AA; 35208 MW; ABBF027E947292FF CRC64;

Query Match 26.1%; Score 97; DB 1; Length 326;
Best Local Similarity 30.5%; Pred. No. 0.00083;
Matches 18; Conservative 7; Mismatches 34; Indels 0; Gaps 0;

Qy 1 FVQPCRDSPPTCGPCPPRHVTFWNYLERCYCNVLCGEREEERACHATHNRACR 59
:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 48 YASRLCGPGSDTVCSPCKNETFTASTNHAPACVSCRGCTGHLSESQCDKTRDRVDC 106
:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

RESULT 8
OX40_HUMAN
ID OX40_HUMAN STANDARD; PRT; 277 AA.
AC P43489; Q13663;
DT 01-NOV-1995 (Rel. 32, Created)
```



```
FT CARBOHYD 52 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 36 C -> Y (IN REF. 2).
FT CONFLICT 173 T -> K (IN REF. 2).
FT CONFLICT 276 N -> S (IN REF. 2).
FT CONFLICT 396 K -> R (IN REF. 2).
SQ SEQUENCE 416 AA; 6BCEAAB54F4D2D56 CRC64;

Query Match 25.7%; Score 95.5; DB 1; Length 416;
Best Local Similarity 33.3%; Pred. No. 0.0015;
Matches 18; Conservative 8; Mismatches 27; Indels 1; Gaps 1;

Oy 5 PCRDSPPTGCPGPPRHYTFQWNY--LERCRCYNVLGGEREEARACHATHNRACR 58
Db 129 PCRDSDTVCECPGTEGDEANFVDPCLPCTI-CEENVVVKECTATSDAECR 181

RESULT 10
OX40_RAT STANDARD; PRT; 271 AA.
AC P15725;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE OX40L RECEPTOR PRECURSOR (OX40 ANTIGEN) (MRC OX40).
GN TNFRSF4 OR TXGPIL OR OX40.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=T-CELL;
RA MEDLINE=90214614; PubMed=2157591;
RT "Characterization of the MRC OX40 antigen of activated CD4 positive T lymphocytes -- a molecule related to nerve growth factor receptor.";
RL EMBO J. 9:1063-1068(1990).
CC -!- FUNCTION: RECEPTOR FOR THE OX40L/GP34 CYTOKINE.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: ACTIVATED T-CELLS.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X17037; CAA34897.1; -
DR PIR; S08036; S08036.
DR PIR; S12783; S12783.
DR HSP; P25942; LCDP.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; 3.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
KW Receptor; T-cell; Antigen; Glycoprotein; Transmembrane; Repeat; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 271 OX40L RECEPTOR.
FT DOMAIN 20 210 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 211 235 POTENTIAL.
FT DOMAIN 236 271 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 25 164 4 X TNFR-CYS.
FT REPEAT 25 60 TNFR-CYS 1.
FT REPEAT 61 102 TNFR-CYS 2.
FT REPEAT 103 123 TNFR-CYS 3 (INCOMPLETE).
FT REPEAT 124 164 TNFR-CYS 4.
FT CARBOHYD 143 143 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 271 AA; 29895 MW; C06465136B16E821 CRC64;
```

```
Query Match 23.7%; Score 88; DB 1; Length 271;
Best Local Similarity 33.3%; Pred. No. 0.0066;
Matches 19; Conservative 6; Mismatches 28; Indels 4; Gaps 2;

Oy 6 CRRDSPPTGCPGPPRHYTFQWNY--LERCRCYNVLGGEREEARACHATHNRACR 60
Db 51 CDHTRDVTCHPCPGFYNEAVNYDTCKQCTQCNRSG--SELKQNCPTPTDVCQR 105

RESULT 11
FASA_PIG STANDARD; PRT; 332 AA.
AC O77736;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE FASL RECEPTOR PRECURSOR (APOPTOSIS-MEDIATING SURFACE ANTIGEN FAS) (APO-1 ANTIGEN) (CD95).
DE TNFRSF6 OR APT1 OR FAS.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN [1]
RP SEQUENCE FROM N.A.
RX Bartling B., Hoffmann J., Holtz J., Schulz R., Heusch G., Darmer D.; "Expression of apoptosis-associated genes in hibernating and stunned myocardium of pig.";
RT myocardiium of pig.;
CC -!- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS FASL. THE ADAPTOR MOLECULE FADD RECRUITS CASPASE-8 TO THE ACTIVATED RECEPTOR. THE RESULTING AGGREGATE CALLED THE DEATH-INDUCING SIGNALING COMPLEX (DISC) PERFORMS CASPASE-8 PROTEOLYTIC ACTIVATION. ACTIVE CASPASE-8 INITIATES THE SUBSEQUENT CASCADE OF CASPASES (ASPARTATE-SPECIFIC CYSTEINE PROTEASES) MEDIATING APOPTOSIS. FAS-MEDIATED APOPTOSIS MAY HAVE A ROLE IN THE INDUCTION OF PERIPHERAL TOLERANCE, IN THE ANTIGEN-STIMULATED SUICIDE OF MATURE T-CELLS, OR BOTH (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- DOMAIN: CONTAINS A DEATH DOMAIN INVOLVED IN THE BINDING OF FADD, AND MAYBE TO OTHER CYTOSOLIC ADAPTOR PROTEINS.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -!- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ001202; CAA04596.1; -
DR INTERPRO; IPR000488; -.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 3.
DR PFAM; PF00531; Death; 1.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
KW Apoptosis; Receptor; Glycoprotein; Transmembrane; Repeat; Signal.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 332 FASL RECEPTOR.
FT DOMAIN 17 175 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 176 192 POTENTIAL.
FT DOMAIN 193 332 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 45 164 3 X TNFR-CYS.
FT REPEAT 45 81 TNFR-CYS 1.
FT REPEAT 82 125 TNFR-CYS 2.
FT REPEAT 126 164 TNFR-CYS 3.
FT DOMAIN 227 311 DEATH DOMAIN.
FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
```

SQ SEQUENCE 332 AA; 37592 MW; 5B8B03682756BFLB CRC64;

Query Match 23.7%; Score 88; DB 1; Length 332;
Best Local Similarity 28.6%; Pred. No. 0.008;
Matches 22; Conservative 7; Mismatches 24; Indels 24; Gaps 3;

QY 8 RDSPTTCGCPP--RH-----YTQFWNYLRCRYCNVLCGERE- 43
DB 52 REQQFCQCPGPKRKHACTSPGAGPQCVPGCEGEDYTKNHHSSKRCRCRVCDGEHL 111

QY 44 EEARACHATHNRACRCR 60
DB 112 EVEKNCRTQNTKCRCK 128

RESULT 12
FASA_RAT
ID FASA_RAT STANDARD; PRT; 324 AA.
AC Q63199;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE FASL RECEPTOR PRECURSOR (APOPTOSIS-MEDIATING SURFACE ANTIGEN FAS)
DE (APO-1 ANTIGEN) (CD95).
GN TNFRSF6 OR PT1 OR FAS.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
SQ SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LIVER;
RX MEDLINE=94128114; PubMed=7507668;
RA Kimura K., Yamamoto M., Wakatsuki T.;
RT "A variant mRNA species encoding a truncated form of Fas antigen in the rat liver.";
RL Biochem. Biophys. Res. Commun. 198:666-674(1994).
CC -1- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS FASL. THE ADAPTOR MOLECULE FADD RECRUITS CASPASE-8 TO THE ACTIVATED RECEPTOR. THE RESULTING AGGREGATE CALLED THE DEATH-INDUCING SIGNALING COMPLEX (DISC) PERFORMS CASPASE-8 PROTEOLYTIC ACTIVATION. ACTIVE CASPASE-8 INITIATES THE SUBSEQUENT CASCADE OF CASPASES (ASPARTATE-SPECIFIC CYSTEINE PROTEASES) MEDIATING APOPTOSIS. FAS-MEDIATED APOPTOSIS MAY HAVE A ROLE IN THE INDUCTION OF PERIPHERAL TOLERANCE, IN THE ANTIGEN-STIMULATED SUICIDE OF MATURE T-CELLS, OR BOTH (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- DOMAIN: CONTAINS A DEATH DOMAIN INVOLVED IN THE BINDING OF FADD, AND MAYBE TO OTHER CYTOSOLIC ADAPTOR PROTEINS.
CC -1- SIMILARITY: CONTAINS A LA-NGR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D26112; BAA05108.1;
CC INTERPRO; IPR000488;
CC INTERPRO; IPR001368;
CC PFAM; PF00020; TNFR_c6; 3.
CC PFAM; PF00531; death; 1.
CC PROSITE; PS00652; TNFR_NGFR_1; 2.
CC PROSITE; PS00050; TNFR_NGFR_2; 2.
CC PROSITE; PS00017; DEATH_DOMAIN; 1.
CC Apoptosis; Receptor; Glycoprotein; Transmembrane; Repeat; Signal.
FT SIGNAL 1 21 BY SIMILARITY.
FT CHAIN 22 324 FASL RECEPTOR.
FT DOMAIN 22 171 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 172 188 POTENTIAL.

FT DOMAIN 189 324 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 43 163 3 X TNFR-CYS.
FT REPEAT 43 79 TNFR-CYS 1.
FT REPEAT 80 123 TNFR-CYS 2.
FT REPEAT 124 163 TNFR-CYS 3.
FT DOMAIN 219 303 DEATH DOMAIN.
FT CARBOHYD 43 43 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 132 132 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 324 AA; 36835 MW; D25D583C909D09D09 CRC64;

Query Match 23.3%; Score 86.5; DB 1; Length 324;
Best Local Similarity 32.4%; Pred. No. 0.011;
Matches 24; Conservative 7; Mismatches 28; Indels 15; Gaps 5;

QY 1 FVORPCR-----RDSPT-----TCGPCPP-RHYTQFWNYLRCRYCNVLCGERE--EEA 46
DB 54 FCCQPCQPGERKVKDCTTSGGAPCTGCTEGEYTDKHYDKKRC-AFCDEGHGLEVE 112

QY 47 RACHATHNRACRCR 60
DB 113 TNCRTQNTKCRCK 126

RESULT 13
FASA_MOUSE
ID FASA_MOUSE STANDARD; PRT; 327 AA.
AC P25446;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE FASL RECEPTOR PRECURSOR (APOPTOSIS-MEDIATING SURFACE ANTIGEN FAS)
DE (APO-1 ANTIGEN) (CD95).
GN TNFRSF6 OR AP1 OR FAS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
SQ SEQUENCE FROM N.A.
RX MEDLINE=92148151; PubMed=1371136;
RA Watanabe-Fukunaga R., Brannan C.I., Itoh N., Yonehara S., Copeland N.G., Jenkins N.A., Nagata S.;
RT "The cDNA structure, expression, and chromosomal assignment of the mouse Fas antigen.";
RL J. Immunol. 148:1274-1279(1992).
RN [2]
SQ SEQUENCE OF 1-96 FROM N.A.
RX MEDLINE=93189576; PubMed=7680478;
RA Adachi M., Watanabe-Fukunaga R., Nagata S.;
RT "Aberrant transcription caused by the insertion of an early transposable element in an intron of the Fas antigen gene of lpr mice.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:1756-1760(1993).
RN [3]
SQ VARIANT LPR.
RX MEDLINE=92195401; PubMed=1372394;
RA Watanabe-Fukunaga R., Brannan C.I., Copeland N.G., Jenkins N.A., Nagata S.;
RT "Lymphoproliferation disorder in mice explained by defects in Fas antigen that mediates apoptosis.";
RL Nature 356:314-317(1992).
CC -1- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS FASL. THE ADAPTOR MOLECULE FADD RECRUITS CASPASE-8 TO THE ACTIVATED RECEPTOR. THE RESULTING AGGREGATE CALLED THE DEATH-INDUCING SIGNALING COMPLEX (DISC) PERFORMS CASPASE-8 PROTEOLYTIC ACTIVATION. ACTIVE CASPASE-8 INITIATES THE SUBSEQUENT CASCADE OF CASPASES (ASPARTATE-SPECIFIC CYSTEINE PROTEASES) MEDIATING APOPTOSIS. FAS-MEDIATED APOPTOSIS MAY HAVE A ROLE IN THE INDUCTION OF PERIPHERAL TOLERANCE, IN THE ANTIGEN-STIMULATED SUICIDE OF MATURE T-CELLS, OR BOTH (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: DETECTED IN VARIOUS TISSUES INCLUDING THYMUS.

```
CC LIVER, LUNG, HEART, AND ADULT OVARY.
CC -!- DOMAIN: CONTAINS A DEATH DOMAIN INVOLVED IN THE BINDING OF FADD,
CC AND MAYBE TO OTHER CYTOSOLIC ADAPTOR PROTEINS.
CC -!- DISEASE: DEFECTS IN FAS ARE THE CAUSE OF A LYMPHOPROLIFERATION
CC DISORDER (LPR) RESPONSIBLE FOR LYMPHADENOPATHY AND AUTOANTIBODY
CC PRODUCTION.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -!- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M83649; AAB37593.1; -
CC EMBL; S56490; AAB25700.1; -
CC EMBL; S56485; AAB25700.1; JOINED.
CC EMBL; S56486; AAB25700.1; JOINED.
CC PIR; A46484; A46484.
CC HSP; P25445; IDDF.
CC MGD; MGI:95484; FAS.
CC INTERPRO; IPR000488; -
CC INTERPRO; IPR001368; -
CC PFAM; PF00020; TNFR_C6; 3.
CC PFAM; PF00531; death; 1.
CC PROSITE; PS00652; TNFR_NGFR_1; 2.
CC PROSITE; PS00650; TNFR_NGFR_2; 2.
CC PROSITE; PS00117; DEATH_DOMAIN; 1.
CC Apoptosis; Receptor; Glycoprotein; Transmembrane; Repeat; Signal;
CC Disease mutation.
CC SIGNAL 1 21
CC CHAIN 22 327 FASL RECEPTOR.
CC DOMAIN 22 169 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 170 186 POTENTIAL.
CC DOMAIN 187 327 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 43 162 3 X TNFR-CYS.
CC REPEAT 43 79 TNFR-CYS 1.
CC REPEAT 80 123 TNFR-CYS 2.
CC REPEAT 124 162 TNFR-CYS 3.
CC DOMAIN 222 306 DEATH DOMAIN.
CC CARBOHYD 43 43 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC VARIANT 246 246 I -> N (IN LPR).
CC SEQUENCE 327 AA; 37418 MW; F0BFFC5ACE3356EEE CRC64;
CC -----
Query Match 23.0%; Score 85.5; DB 1; Length 327;
Best Local Similarity 33.9%; Pred. No. 0.015;
Matches 20; Conservative 8; Mismatches 26; Indels 5; Gaps 4;
Qy 6 CRRDPT-TCGCP-PPHRYTQFNWY--EERACHATHNACRCR 60
| : | | | | : | : | | | | | | | | | | | | |
Db 69 CKMGTTTCAPCTEGKEYMDKNHYADKRCR-TLDCDEHGLEVTNCTLTQNTKCK 126
| : | | | | : | : | | | | | | | | | | | | |
RESULT 14
OX40_MOUSE 272 AA; PRT; 272 AA.
AC P47741:
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE OX40L RECEPTOR PRECURSOR (OX40 ANTIGEN).
GN TNFRS4 OR TXGP1 OR OX40.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
```

```
RX MEDLINE-94044750; PubMed-8228223;
RA Calderhead D.M., Buhlmann J.E., van den Eertwegh A.J.,
RA Claassen E., Noelle R.J., Fell H.;
RT "Cloning of mouse OX40: a T cell activation marker that may mediate
RT T-B cell interactions.";
RL J. Immunol. 151:5261-5271(1993).
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE-95255413; PubMed-7737295;
RA Birkeland M.L., Copeland N.G., Gilbert D.J., Jenkins N.A.,
RA Barclay A.N.;
RT "Gene structure and chromosomal localization of the mouse homologue
RT of rat OX40 protein.";
RL Eur. J. Immunol. 25:926-930(1995).
CC -!- FUNCTION: RECEPTOR FOR THE OX40L/GP34 CYTOKINE.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z21674; CAA79772.1; -
CC EMBL; X85214; CAA59476.1; -
CC HSP; P25942; ICDF.
CC MGD; MGI:104512; TXGP1.
CC INTERPRO; IPR001368; -
CC PFAM; PF00020; TNFR_C6; 3.
CC PROSITE; PS00652; TNFR_NGFR_1; 3.
CC PROSITE; PS00650; TNFR_NGFR_2; 2.
CC Receptor; T-cell; Antigen; Glycoprotein; Transmembrane; Repeat;
CC Signal.
CC SIGNAL 1 19 POTENTIAL.
CC CHAIN 20 272 OX40L RECEPTOR.
CC DOMAIN 20 211 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 212 236 POTENTIAL.
CC DOMAIN 237 272 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 26 165 4 X TNFR-CYS.
CC REPEAT 26 61 TNFR-CYS 1.
CC REPEAT 62 103 TNFR-CYS 2.
CC REPEAT 104 124 TNFR-CYS 3 (INCOMPLETE).
CC REPEAT 125 165 TNFR-CYS 4.
CC CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CONFLICT 15 15 A -> G (IN REF. 2).
CC SEQUENCE 272 AA; 30153 MW; 06E7BB4156F0D08E CRC64;
CC -----
Query Match 22.4%; Score 83; DB 1; Length 272;
Best Local Similarity 33.3%; Pred. No. 0.023;
Matches 19; Conservative 5; Mismatches 29; Indels 4; Gaps 2;
Qy 6 CRRDPTTCGCP-PPHRYTQFNWY--LERCRCNVLCGEREEARACHATHNACRCR 60
| : | | | | : | : | | | | | | | | | | | | |
Db 52 CDHTRDTLCHPCETGTFYNEAVNYDTCKQCTQCNRHSG--SELKQNCPTPTQDTCVCR 106
| : | | | | : | : | | | | | | | | | | | | |
RESULT 15
TNRI_BOVIN
ID TNRI_BOVIN STANDARD; PRT; 471 AA.
AC O19131;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR 1 PRECURSOR (P60) (TNF-RI)
DE (P55).
GN TNFRSF1A OR TNFR1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
```

OC Bovidae; Bovinae; Bos.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=AORTA;
RA Lee E.-K., Taylor M.J., Kehrli M.E.;
RT "Cloning of cDNA encoding bovine tumor necrosis factor-receptor I
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: RECEPTOR FOR TNF-ALPHA. THE ADAPTOR MOLECULE FADD
CC RECRUITS CASPASE-8 TO THE ACTIVATED RECEPTOR. THE RESULTING
CC AGGREGATE CALLED THE DEATH-INDUCING SIGNALING COMPLEX (DISC)
CC PERFORMS CASPASE-8 PROTEOLYTIC ACTIVATION WHICH INITIATES THE
CC SUBSEQUENT CASCADE OF CASPASES (ASPARTATE-SPECIFIC CYSTEINE
CC PROTEASES) MEDIATING APOPTOSIS (BY SIMILARITY).
CC -1- SUBUNIT: TNF BINDING TO THE EXTRACELLULAR DOMAIN OF TNFR1 LEADS TO
CC HOMOTRIMERIZATION. ONCE AGGREGATED THE RECEPTORS DEATH DOMAINS
CC PROVIDE A NOVEL MOLECULAR INTERFACE THAT INTERACTS SPECIFICALLY
CC WITH THE DEATH DOMAIN OF TRADD. VARIOUS TRADD-INTERACTING
CC PROTEINS SUCH AS TRAFs, RIP AND POSSIBLY FADD, ARE RECRUITED TO
CC TNFR1 COMPLEX BY THEIR ASSOCIATION WITH TRADD. THIS COMPLEX
CC ACTIVATES AT LEAST TWO DISTINCT SIGNALING CASCADES, APOPTOSIS AND
CC NF-KAPPA B SIGNALING (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; U90937; AAB65143.1; -
DR HSSP; P19438; ITNR.
DR INTERPRO; IPR000488; -
DR INTERPRO; IPR001368; -
DR PFAM; PF00020; TNFR_c6; 3.
DR PFAM; PF00531; death; 1.
DR PROSITE; PS00652; TNFR_NGFR_1; 3.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal; Apoptosis.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 471 TUMOR NECROSIS FACTOR RECEPTOR 1.
FT DOMAIN 22 210 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 211 233 POTENTIAL.
FT DOMAIN 234 471 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 43 195 4 X TNFR-CYS.
FT REPEAT 43 82 TNFR-CYS 1.
FT REPEAT 83 125 TNFR-CYS 2.
FT REPEAT 126 166 TNFR-CYS 3.
FT REPEAT 167 195 TNFR-CYS 4.
FT DOMAIN 340 360 N-SMASE ACTIVATION DOMAIN (NSD).
FT DOMAIN 372 457 DEATH DOMAIN.
FT DISULFID 44 58 BY SIMILARITY.
FT DISULFID 59 72 BY SIMILARITY.
FT DISULFID 62 81 BY SIMILARITY.
FT DISULFID 84 99 BY SIMILARITY.
FT DISULFID 102 117 BY SIMILARITY.
FT DISULFID 105 125 BY SIMILARITY.
FT DISULFID 127 143 BY SIMILARITY.
FT DISULFID 146 158 BY SIMILARITY.
FT DISULFID 149 166 BY SIMILARITY.
FT DISULFID 168 179 BY SIMILARITY.
FT DISULFID 182 190 BY SIMILARITY.
FT DISULFID 185 194 BY SIMILARITY.
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 471 AA; 51367 MW; 5243EF514DFE81C4 CRC64;

Query Match 21.6%; Score 80; DB 1; Length 471;
Best Local Similarity 26.8%; Pred. No. 0.081;
Matches 22; Conservative 8; Mismatches 26; Indels 26; Gaps 5;
QY 5 PCRRDSPTTCGCPPRHYTQFWNYLE-RCRYC-----NVLCGEREB----- 44
DB 116 PCVVDRTVCG-CRKNQVREYWGTFGRCLNCSLCPNGTNTNIPCOERQDTICHMGEFL 174
QY 45 ---EARACHATHNRACR--CRT 61
DB 175 KGAKCISCHDCRKECEKLCPT 196

Search completed: January 30, 2001, 17:05:56
Job time: 1217 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:54:41 ; Search time 254.27 Seconds
(without alignments)
28.118 Million cell updates/sec

Title: US-09-518-931-2_COPY_57_117

Perfect score: 371

Sequence: 1 FVQPCRDSPITGCPGPPR.....REEERACHATHNRCRCRT 61

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_15:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	371	100.0	300	4	O95407
2	147	39.6	372	4	Q9UHP4
3	147	39.6	401	4	Q00300
4	138	37.2	401	11	O08712
5	135	36.4	401	11	O08727
6	130	35.0	348	12	O57277
7	130	35.0	348	12	O57103
8	130	35.0	348	12	O57108
9	130	35.0	349	12	O57291
10	130	35.0	349	12	O57099
11	130	35.0	349	12	O57100
12	130	35.0	349	12	O57101
13	130	35.0	349	12	O57102
14	129	34.8	459	11	Q62327
15	128	34.5	439	4	Q16042
16	126	34.0	348	12	O57112
17	126	34.0	348	12	Q85407
18	126	34.0	349	12	O57109
19	126	34.0	349	12	O57110

20	126	34.0	349	12	O57111	057111 variola vir
21	126	34.0	349	12	O89118	089118 variola vir
22	126	34.0	349	12	O89098	089098 variola vir
23	125	33.7	349	12	O57284	057284 camelpox vi
24	125	33.7	349	12	O57098	057098 camelpox vi
25	124	33.4	326	12	O57120	057120 cowpox viru
26	122.5	33.0	302	13	O9PU50	O9PU50 salvelinus
27	122.5	33.0	482	11	O88734	O88734 mus musculu
28	121	32.6	349	12	O57097	O57097 camelpox vi
29	120	32.3	326	12	O57122	O57122 cowpox viru
30	120	32.3	351	12	O57117	O57117 cowpox viru
31	120	32.3	351	12	O73559	O73559 cowpox viru
32	119	32.1	355	12	O85308	O85308 cowpox viru
33	118	31.8	350	12	O57116	O57116 cowpox viru
34	118	31.8	350	12	O57123	O57123 cowpox viru
35	114	30.7	349	12	O57305	O57305 cowpox viru
36	112	30.2	347	12	O57119	O57119 cowpox viru
37	112	30.2	351	12	O57121	O57121 cowpox viru
38	112	30.2	360	12	O57118	O57118 cowpox viru
39	110	29.6	347	12	O57115	O57115 cowpox viru
40	107	28.8	167	12	O72762	O72762 cowpox viru
41	103	27.8	322	12	O72761	O72761 cowpox viru
42	101	27.2	320	12	O57079	O57079 cowpox viru
43	99	26.7	267	6	O02764	O02764 oryctolagus
44	99	26.7	316	12	O57092	O57092 ectromelia
45	99	26.7	320	12	O57091	O57091 ectromelia

ALIGNMENTS

RESULT 1

O95407
ID O95407 PRELIMINARY; PRT; 300 AA.
AC O95407;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE DECOY RECEPTOR 3 (M68) (M68C) (M68E).
GN DCR3 OR TR6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99087326; PubMed=9872321;
RA Pitti R.M., Marsters S.A., Lawrence D.A., Roy M., Kischkel F.C.,
RA Dowd P., Huang A., Donahue C.J., Sherwood S.W., Baldwin D.T.,
RA Godowski P.J., Wood W.I., Gurney A.L., Hillan K.J., Cohen R.L.,
RA Goddard A.D., Botstein D., Ashkenazi A.;
RT "Genomic amplification of a decoy receptor for Fas ligand in lung and
RT colon cancer.";
RL Nature 396:699-703(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Blood;
RA Yu K.Y., Kwon B., Ni J., Zhai Y., Ebner R., Kwon B.S.;
RT "A newly identified member of tumor necrosis factor receptor
RT superfamily (TR6) suppresses LIGHT-mediated apoptosis.";
RL J. Biol. Chem. 274:13733-13736(1999).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=pancreas;
RA MEDLINE=2012600; PubMed=10655513;
RA Bai C., Connolly B., Metzker M.L., Hilliard C.A., Liu X., Sandig V.,
RT Soderman A., Galloway S.M., Liu Q., Austin C.P., Caskey C.T.;
RT "Overexpression of M68/Dcr3 in human gastrointestinal tract tumors
RT independent of gene amplification and its location in a four-gene
RL cluster.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:1230-1235(2000).
DR EMBL; AF104419; AAD03056.1; -.

```

DR EMBL: AF134240; RAD29688.1; -
DR EMBL: AF217796; AAF33244.1; -
DR EMBL: AF217793; AAF33685.1; -
DR EMBL: AF217794; AAF33686.1; -
DR HSP: P25942; ICDF.
DR INTERPRO: IPR000561; -
DR INTERPRO: IPR001368; -
DR PFAM: PF00020; TNFR_C6; 4.
DR PROSITE: PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
DR PROSITE: PS50050; TNFR_NGFR_2; 2.
DR PRODOM: PD000771; -; 1.
KW Receptor.
SQ SEQUENCE 300 AA; 32679 MW; F90AEE33718449AF CRC64;

Query Match 100.0%; Score 371; DB 4; Length 300;
Best Local Similarity 100.0%; Pred. No. 1.2e-36;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVQPCRDSPPTCGPCPPRHYYTQFNWYLCRYCNVLCGEREEARACHATHNRACRCR 60
Db 57 FVQPCRDSPPTCGPCPPRHYYTQFNWYLCRYCNVLCGEREEARACHATHNRACRCR 116

QY 61 T 61
Db 117 T 117

RESULT 2
Q9UHP4 ID Q9UHP4 PRELIMINARY; PRT; 372 AA.
AC Q9UHP4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE OSTEOPROTEGERIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=PLACENTA;
RC He Z.-Y., Yang G.-Z., Zhang W.-J., Wu X.-F.;
RT "Cloning and Expression of Osteoprotegerin from Homo sapiens.";
RL Sheng Wu Hua Hsueh Yu Sheng Wu Wu Li Hsueh Pao 31:680-684(1999).
DR EMBL: AF134187; AAF20168.1; -
DR HSP: P25942; ICDF.
DR INTERPRO: IPR001368; -
DR PFAM: PF00020; TNFR_C6; 3.
DR PROSITE: PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE: PS50050; TNFR_NGFR_2; 2.
DR NON_TER 1
FT SEQUENCE 372 AA; 42758 MW; F02527A5CD01CCD3 CRC64;

Query Match 39.68; Score 147; DB 4; Length 372;
Best Local Similarity 35.08; Pred. No. 5.9e-10;
Matches 21; Conservative 13; Mismatches 26; Indels 0; Gaps 0;

QY 1 FVQPCRDSPPTCGPCPPRHYYTQFNWYLCRYCNVLCGEREEARACHATHNRACRCR 60
Db 28 YLKQHTAKWTKVACPCPDHYTDSWHTSDCLYCSPVKELQYVKBNCNTHNRVCECK 87

RESULT 3
ID Q00300 PRELIMINARY; PRT; 401 AA.
AC Q00300; O60236;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)

```

```

DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)
DE (OCIF) (TUMOR NECROSIS FACTOR RECEPTOR SUPERFAMILY MEMBER 11b).
CN TNFRSF11B OR OPG OR OCIF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=KIDNEY.
RC MEDLINE=97262071; PubMed=9108485;
RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,
RA Luethy R., Nguyen H.Q., Wooden S., Bennett L., Boone T., Shimamoto G.,
RA Derose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,
RA Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,
RA Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,
RA Sugis S., Boyle W.J.;
RT "Osteoprotegerin: a novel secreted protein involved in the regulation
RT of bone density.";
RL Cell 89:309-319(1997).
[2]
RN SEQUENCE FROM N.A.
RP TISSUE=LUNG FIBROBLAST;
RC MEDLINE=98151033; PubMed=9492069;
RA Yasuda H., Shima N., Nakagawa N., Mochizuki S.-I., Yano K., Fujise N.,
RA Sato Y., Goto M., Yamaguchi K., Kuriyama M., Kanno T., Murakami A.,
RA Tsuda E., Morinaga T., Higashio K.;
RT "Identity of osteoclastogenesis inhibitory factor (OCIF) and
RT osteoprotegerin (OPG): a mechanism by which OPG/OCIF inhibits
RT osteoclastogenesis in vitro.";
RL Endocrinology 139:1329-1337(1998).
[3]
RN SEQUENCE FROM N.A.
RP TISSUE=PLACENTA;
RC MEDLINE=983551569; PubMed=9688283;
RA Morinaga T., Nakagawa N., Yasuda H., Tsuda E., Higashio K.;
RT "Cloning and characterization of the gene encoding human
RT osteoprotegerin/osteoclastogenesis-inhibitory factor.";
RL Eur. J. Biochem. 254:685-691(1998).
CC -!- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES
CC OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY
CC SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO
CC OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN
CC STROMAL CELLS AND OSTEOCLAST PROGENITORS.
CC -!- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM) (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGHEST LEVELS IN LUNG, HEART,
CC KIDNEY, PLACENTA, THYROID, SPINAL CORD AND LIVER. ALSO DETECTED IN
CC A NUMBER OF OTHER HEMATOPOIETIC AND IMMUNE ORGANS. NOT DETECTED IN
CC THE PANCREAS OR PERIPHERAL BLOOD LYMPHOCYTES.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC EMBL: AB002146; BAA25910.1; -
CC EMBL: AB008822; BAA32076.1; -
CC EMBL: AB008821; BAA32076.1; JOINED.
CC EMBL: U94332; AAB3709.1; -
CC HSP: P25942; ICDF.
CC MIM: 602643; -
CC INTERPRO: IPR001368; -
CC PFAM: PF00020; TNFR_C6; 3.
CC PROSITE: PS00652; TNFR_NGFR_1; UNKNOWN_1.
CC PROSITE: PS50050; TNFR_NGFR_2; 2.
CC PRODOM: PD000771; -; 1.
KW Glycoprotein; Repeat; Cytokine; Signal.
FT SIGNAL 1 21 BY SIMILARITY.
FT CHAIN 22 401 OSTEOPROTEGERIN.
FT DOMAIN 23 183 4 X TNFR-CYS.
FT REPEAT 23 63 TNFR-CYS 1.
FT REPEAT 64 106 TNFR-CYS 2.
FT REPEAT 107 143 TNFR-CYS 3.
FT REPEAT 144 201 TNFR-CYS 4.
FT DOMAIN 306 365 DEATH DOMAIN.
FT DISULFID 41 54 BY SIMILARITY.

```

```

FT DISULFID 44 62 BY SIMILARITY.
FT DISULFID 65 80 BY SIMILARITY.
FT DISULFID 83 97 BY SIMILARITY.
FT DISULFID 87 105 BY SIMILARITY.
FT DISULFID 118 142 BY SIMILARITY.
FT DISULFID 145 160 BY SIMILARITY.
FT CARBOHYD 98 98 POTENTIAL.
FT CARBOHYD 152 152 POTENTIAL.
FT CARBOHYD 165 165 POTENTIAL.
FT CARBOHYD 178 178 POTENTIAL.
FT CARBOHYD 289 289 POTENTIAL.
FT CARBOHYD 263 263 A -> D (IN REF. 2 AND 3).
SQ SEQUENCE 401 AA; 45996 MW; EB42FA51C9D7C71E CRC64;

Query Match 39.6%; Score 147; DB 4; Length 401;
Best Local Similarity 35.0%; Pred. No. 6.3e-10;
Matches 21; Conservative 13; Mismatches 26; Indels 0; Gaps 0;

Qy 1 FVQPCRDSPPTGCPGPPRHVYTFWNYLCRCYCNVLCGEREEARACHATHNRACRCR 60
Db 49 YLKQHCATKWTVCAPCPDHYTDSWHTSDCLYCSPVKELQYVKQECNTHNRVCECK 108

RESULT 4
O08712 PRELIMINARY; PRT; 401 AA.
AC O08712; 070202;
DT 01-JUL-1997 (Tremblrel. 04, Created)
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)
DE (OCIF).
GN TNFRSF11B OR OPG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=KIDNEY;
RX MEDLINE=97262071; PubMed=9108485;
RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,
RA Luethy R., Nguyen H.Q., Wooden S., Bennett L., Boone T., Shimamoto G.,
RA Derose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,
RA Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,
RA Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,
RA Suggs S., Boyle W.J.;
RT "Osteoprotegerin: a novel secreted protein involved in the regulation
RT of bone density."
RL Cell 89:309-319(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/OLA, AND NIH SWISS;
RX MEDLINE=98382527; PubMed=9714833;
RA Mizuno A., Murakami A., Nakagawa N., Yasuda H., Tsuda E., Morinaga T.,
RA Higashio K.;
RT "Structure of the mouse osteoclastogenesis inhibitory factor (OCIF)
RT gene and its expression in embryogenesis."
RL Gene 215:339-343(1998).
CC -!- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES
CC OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY
CC SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO
CC OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN
CC STROMAL CELL AND OSTEOCLAST PROGENITORS.
CC -!- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM).
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN LIVER, LUNG,
CC BRAIN, HEART, KIDNEY, STOMACH, INTESTINE, SKIN, CALVARIA AND
CC PLACENTA. NOT DETECTED IN SPLEEN.
CC -!- DEVELOPMENTAL STAGE: IN THE EMBRYO, HIGH LEVELS ARE DETECTED AT
CC DAY 7. AT DAY 11, EXPRESSION DECREASES AND THEN INCREASES FROM DAY
CC 15 TO DAY 17.

```

```

CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL; U94331; AAB53708.1; -.
DR EMBL; AB013898; BAA28269.1; -.
DR EMBL; AB013903; BAA33388.1; -.
DR EMBL; AB013899; BAA33388.1; JOINED.
DR EMBL; AB013900; BAA33388.1; JOINED.
DR EMBL; AB013901; BAA33388.1; JOINED.
DR EMBL; AB013902; BAA33388.1; JOINED.
DR HSP; P25942; ICDF.
DR MGD; MGI:109587; Opg.
DR INTERPRO; IPR000488; -.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS00117; DEATH DOMAIN; 1.
DR PROSITE; PS00050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -. 1.
KW Glycoprotein; Repeat; Cytokine; Signal.
FT SIGNAL 1 21 OSTEOPROTEGERIN.
FT CHAIN 22 401 4 X TNFR-CYS.
FT DOMAIN 23 201 TNFR-CYS 1.
FT REPEAT 23 63 TNFR-CYS 2.
FT REPEAT 64 106 TNFR-CYS 3.
FT REPEAT 107 143 TNFR-CYS 4.
FT REPEAT 144 201 DEATH DOMAIN.
FT DOMAIN 306 365 BY SIMILARITY.
FT DISULFID 41 54 BY SIMILARITY.
FT DISULFID 44 62 BY SIMILARITY.
FT DISULFID 65 80 BY SIMILARITY.
FT DISULFID 83 97 BY SIMILARITY.
FT DISULFID 87 105 BY SIMILARITY.
FT DISULFID 118 142 BY SIMILARITY.
FT DISULFID 145 160 BY SIMILARITY.
FT CARBOHYD 98 98 POTENTIAL.
FT CARBOHYD 165 165 POTENTIAL.
FT CARBOHYD 178 178 POTENTIAL.
FT CARBOHYD 289 289 POTENTIAL.
FT VARIANT 138 138 R -> P (IN STRAINS 129/OLA AND NIH SWISS).
FT VARIANT 161 161 I -> R (IN STRAINS 129/OLA AND NIH SWISS).
FT VARIANT 165 165 N -> D (IN STRAINS 129/OLA AND NIH SWISS).
FT VARIANT 288 288 S -> A (IN STRAINS 129/OLA AND NIH SWISS).
FT VARIANT 296 296 L -> R (IN STRAINS 129/OLA AND NIH SWISS).
SQ SEQUENCE 401 AA; 45923 MW; CAA6102D3B312470 CRC64;

Query Match 37.2%; Score 138; DB 11; Length 401;
Best Local Similarity 35.6%; Pred. No. 7.4e-09;
Matches 21; Conservative 11; Mismatches 27; Indels 0; Gaps 0;

Qy 1 FVQPCRDSPPTGCPGPPRHVYTFWNYLCRCYCNVLCGEREEARACHATHNRACRCR 59
Db 49 YLKQHCATKWTVCAPCPDHYTDSWHTSDCLYCSPVKELQYVKQECNTHNRVCECK 107

RESULT 5
O08727 PRELIMINARY; PRT; 401 AA.
ID O08727
AC O08727;
DT 01-JUL-1997 (Tremblrel. 04, Created)
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)
DE (OCIF).
GN TNFRSF11B OR OPG.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;

```



```
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10244;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZAIRE-1970 (CONGO-8);
RA Loparev V.N., Parsons J.M., Esposito J.J.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U88142; AAB94367.1; -.
DR HSSP; P25942; ICDF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 2.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -.
DR PRODOM; PD000771; -.
SQ SEQUENCE 348 AA; 38212 MW; E555979057DEC91F CRC64;

Query Match          35.0%; Score 130; DB 12; Length 348;
Best Local Similarity 37.3%; Pred. No. 5.9e-08;
Matches 22; Conservative 8; Mismatches 29; Indels 0; Gaps 0;

Qy 1 FVQPCRRDSTTCGCPPRHYTQFWNYLRCRYCNVLCGEREEARACHATHNRACRC 59
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 52 YASRLCDSKTTQCTPCGSDTFTSHNNHLQACLSNGRCDSNQVETRSCNTTHNRICEC 110

RESULT 9
ID O57291 PRELIMINARY; PRT; 349 AA.
AC O57291;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN CRMB.
OS Monkeypox virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10244;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VARIOUS STRAINS;
RA Loparev V.N., Parsons J.M., Esposito J.J.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; U88144; AAB94369.1; -.
DR EMBL; U87842; AAB94359.1; -.
DR EMBL; U87994; AAB94365.1; -.
DR EMBL; U87995; AAB94366.1; -.
DR EMBL; U88143; AAB94368.1; -.
DR HSSP; P25942; ICDF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 2.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -.
DR PRODOM; PD000771; -.
SQ SEQUENCE 349 AA; 38295 MW; CBD2C949ED2B8E7C CRC64;

Query Match          35.0%; Score 130; DB 12; Length 349;
Best Local Similarity 37.3%; Pred. No. 5.9e-08;
Matches 22; Conservative 8; Mismatches 29; Indels 0; Gaps 0;

Qy 1 FVQPCRRDSTTCGCPPRHYTQFWNYLRCRYCNVLCGEREEARACHATHNRACRC 59
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 52 YASRLCDSKTTQCTPCGSDTFTSHNNHLQACLSNGRCDSNQVETRSCNTTHNRICEC 110

RESULT 10
O57099 PRELIMINARY; PRT; 349 AA.
ID O57099
AC O57099;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
```

```
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN CRMB.
OS Monkeypox virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10244;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SIERRA LEONE-1970 (70-0266);
RA Loparev V.N., Parsons J.M., Esposito J.J.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U87843; AAB94360.1; -.
DR HSSP; P25942; ICDF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 2.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -.
DR PRODOM; PD000771; -.
SQ SEQUENCE 349 AA; 38321 MW; FE449028CC933F57 CRC64;

Query Match          35.0%; Score 130; DB 12; Length 349;
Best Local Similarity 37.3%; Pred. No. 5.9e-08;
Matches 22; Conservative 8; Mismatches 29; Indels 0; Gaps 0;

Qy 1 FVQPCRRDSTTCGCPPRHYTQFWNYLRCRYCNVLCGEREEARACHATHNRACRC 59
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 52 YASRLCDSKTTQCTPCGSDTFTSHNNHLQACLSNGRCDSNQVETRSCNTTHNRICEC 110

RESULT 11
ID O57100 PRELIMINARY; PRT; 349 AA.
AC O57100;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN CRMB.
OS Monkeypox virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10244;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIGERIA-1971 (71-0082);
RA Loparev V.N., Parsons J.M., Esposito J.J.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U87844; AAB94361.1; -.
DR HSSP; P25942; ICDF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 2.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -.
DR PRODOM; PD000771; -.
SQ SEQUENCE 349 AA; 38239 MW; DF6C280D478F2422 CRC64;

Query Match          35.0%; Score 130; DB 12; Length 349;
Best Local Similarity 37.3%; Pred. No. 5.9e-08;
Matches 22; Conservative 8; Mismatches 29; Indels 0; Gaps 0;

Qy 1 FVQPCRRDSTTCGCPPRHYTQFWNYLRCRYCNVLCGEREEARACHATHNRACRC 59
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 52 YASRLCDSKTTQCTPCGSDTFTSHNNHLQACLSNGRCDSNQVETRSCNTTHNRICEC 110

RESULT 12
O57101 PRELIMINARY; PRT; 349 AA.
ID O57101
AC O57101;
```

[illegible]

Matches 21; Conservative 5; Mismatches 29; Indels 0; Gaps 0;

QY 6 CRRDSPTCGCPPRHYTQFWNYLERCRYCNVLCGEREEEARACHATHNRACRCR 60

Db 45 CTKTSDTVCDSCEDSTYTLWNVPECLSCGSRCSDDQVETQACTREQNRICR 99

Search completed: January 30, 2001, 16:54:41
Job time: 897 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: January 30, 2001, 16:47:41 ; Search time 132.77 Seconds
(without alignments)
8.250 Million cell updates/sec

Title: US-09-518-931-2_COPY_57_117

Perfect score: 371

Sequence: 1 FVQPCRRDPTTCGPCPPR.....REEEARACHATHNACRCRT 61

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6_COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	371	100.0	300	2	US-08-794-796-2
2	147	39.6	401	3	US-08-974-022-6
3	138	37.2	401	3	US-08-974-022-4
4	135	36.4	401	3	US-08-974-022-2
5	130	35.0	163	2	US-08-219-237B-5
6	128	34.5	120	3	US-08-974-022-42
7	128	34.5	164	2	US-08-232-087A-9
8	128	34.5	227	3	US-08-974-022-48
9	128	34.5	461	1	US-08-385-229-2
10	128	34.5	461	5	US-08-650-000-2
11	128	34.5	461	5	5395760-2
12	128	34.5	486	1	US-08-243-010-1
13	128	34.0	518	1	US-08-385-229-4
14	126	34.0	474	2	US-08-650-000-4
15	126	34.0	474	5	5395760-4
16	119	32.1	355	1	US-08-292-549-6
17	106	28.6	207	3	US-08-974-022-47
18	106	28.6	325	1	US-08-292-549-2
19	106	28.6	325	4	PCT-US91-02207-2
20	103.5	27.9	197	2	US-08-505-606-1
21	97	26.1	326	1	US-08-292-549-4
22	97	26.1	326	4	PCT-US91-02207-4
23	96	25.9	277	2	US-08-147-784-2
24	91	24.5	42	1	US-08-050-319B-32
25	91	24.5	42	2	US-08-465-382-32
26	90.5	24.4	283	4	PCT-US96-12374-2
27	88	23.7	139	2	US-08-219-237B-8
28	88	23.7	205	3	US-08-974-022-51

29	87.5	23.6	440	3	US-08-883-036A-2	Sequence 2, Appli
30	86.5	23.3	186	1	US-08-089-458B-6	Sequence 6, Appli
31	86	23.2	655	3	US-08-959-382-2	Sequence 2, Appli
32	83	22.4	206	1	US-08-097-827-7	Sequence 7, Appli
33	83	22.4	206	1	US-08-494-574-7	Sequence 7, Appli
34	83	22.4	438	1	US-08-097-827-11	Sequence 11, Appl
35	83	22.4	438	1	US-08-494-574-11	Sequence 11, Appl
36	79.5	21.4	119	2	US-08-219-237B-3	Sequence 3, Appli
37	79.5	21.4	219	3	US-08-974-022-45	Sequence 45, Appl
38	79.5	21.4	314	1	US-08-444-231-19	Sequence 19, Appl
39	79.5	21.4	314	1	US-08-152-443A-19	Sequence 19, Appl
40	79.5	21.4	314	4	PCT-US95-17083-4	Sequence 4, Appli
41	79.5	21.4	335	2	US-08-219-237B-2	Sequence 2, Appli
42	79.5	21.4	335	2	US-08-409-338-1	Sequence 1, Appli
43	79.5	21.4	335	3	US-08-815-469-6	Sequence 6, Appli
44	79.5	21.4	335	4	PCT-US95-17083-2	Sequence 2, Appli
45	78.5	21.2	191	3	US-08-974-022-52	Sequence 52, Appl

ALIGNMENTS

RESULT 1
US-08-794-796-2
; Sequence 2, Application US/08794796
; Patent No. 5885800
; GENERAL INFORMATION:
; APPLICANT: Emery, John
; APPLICANT: Tan, KB
; APPLICANT: Truneh, Alem
; APPLICANT: Young, Peter
; TITLE OF INVENTION: Tumor Necrosis Related Receptor,
; TITLE OF INVENTION: TR4
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithkline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/794,796
FILING DATE: 04-FEB-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Han, William T
REGISTRATION NUMBER: 34,344
REFERENCE/DOCKET NUMBER: GH50000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5219
TELEFAX: 610-270-4026
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 300 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-794-796-2

Query Match 100.0%; Score 371; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 2e-32;

```
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FVORCRDRSPPTCGCPRRHYTQFWNYLRCRYCNVLCGEREEARACHATHNRACRCR 60
Db 57 FVORCRDRSPPTCGCPRRHYTQFWNYLRCRYCNVLCGEREEARACHATHNRACRCR 116
QY 61 T 61
Db 117 T 117

RESULT 2
US-08-974-022-6
; Sequence 6, Application US/08974022
; Patent No. 6015938
; GENERAL INFORMATION:
; APPLICANT: Boyle, William J.
; APPLICANT: Lacey, David L.
; APPLICANT: Calzone, Frank J.
; APPLICANT: Chang, Ming-Shi
; TITLE OF INVENTION: OSTEOPROTEGERIN
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,022
; FILING DATE: 12-DEC-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/577,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-378
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 401 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-974-022-6
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,022
; FILING DATE: 12-DEC-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/577,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-378
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 401 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-974-022-6

Query Match 39.6%; Score 147; DB 3; Length 401;
Best Local Similarity 35.0%; Pred. No. 1.6e-08;
Matches 21; Conservative 13; Mismatches 26; Indels 0; Gaps 0;
QY 1 FVORCRDRSPPTCGCPRRHYTQFWNYLRCRYCNVLCGEREEARACHATHNRACRCR 60
Db 49 YLKQHTAKWKTVACPDHYTDSWHTSDCLYCSVCKELQYVKQECNRTHNRVCECK 108

RESULT 3
US-08-974-022-4
; Sequence 4, Application US/08974022
; Patent No. 6015938
; GENERAL INFORMATION:
; APPLICANT: Boyle, William J.
; APPLICANT: Lacey, David L.
; APPLICANT: Calzone, Frank J.
; APPLICANT: Chang, Ming-Shi
; TITLE OF INVENTION: OSTEOPROTEGERIN
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,022
; FILING DATE: 12-DEC-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/577,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-378
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 401 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-974-022-4

Query Match 37.2%; Score 138; DB 3; Length 401;
Best Local Similarity 35.6%; Pred. No. 1.4e-07;
Matches 21; Conservative 11; Mismatches 27; Indels 0; Gaps 0;
QY 1 FVORCRDRSPPTCGCPRRHYTQFWNYLRCRYCNVLCGEREEARACHATHNRACRCR 59
Db 49 YLKQHTVRRKTLVCPDHSYTDWHTSDCLYCSVCKELQYVKQECNRTHNRVCECK 107

RESULT 4
US-08-974-022-2
; Sequence 2, Application US/08974022
; Patent No. 6015938
; GENERAL INFORMATION:
; APPLICANT: Boyle, William J.
; APPLICANT: Lacey, David L.
; APPLICANT: Calzone, Frank J.
; APPLICANT: Chang, Ming-Shi
; TITLE OF INVENTION: OSTEOPROTEGERIN
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,022
; FILING DATE: 12-DEC-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/577,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-378
```



```
; FILING DATE: 08-SEP-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 756-103P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 164 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..164
; OTHER INFORMATION: /note= "TNFR2, see Fig. 5"
; US-08-232-087A-9

Query Match 34.5%; Score 128; DB 2; Length 164;
Best Local Similarity 38.2%; Pred. No. 6.9e-07;
Matches 21; Conservative 5; Mismatches 29; Indels 0; Gaps 0;

QY 6 CRRDSPTCGCPPRHYTQFWNYLRCRYCNVLCGEREEERACHATHNRCRCR 60
Db 30 CTKTSDTVCDSCEDSYTOLMNNVPECLSGRSCSSDQVETQACTREQNRICR 84

RESULT 8
US-08-974-022-48
; Sequence 48, Application US/08974022
; Patent No. 6015938
; GENERAL INFORMATION:
; APPLICANT: Boyle, William J.
; APPLICANT: Lacey, David L.
; APPLICANT: Calzone, Frank J.
; APPLICANT: Chang, Ming-Shi
; TITLE OF INVENTION: OSTEOPROTEGERIN
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 Behavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,022
; FILING DATE: 12-DEC-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/577,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-378
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 227 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

; FILING DATE: 08-SEP-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 756-103P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 164 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..164
; OTHER INFORMATION: /note= "TNFR2, see Fig. 5"
; US-08-232-087A-9

Query Match 34.5%; Score 128; DB 2; Length 164;
Best Local Similarity 38.2%; Pred. No. 6.9e-07;
Matches 21; Conservative 5; Mismatches 29; Indels 0; Gaps 0;

QY 6 CRRDSPTCGCPPRHYTQFWNYLRCRYCNVLCGEREEERACHATHNRCRCR 60
Db 30 CTKTSDTVCDSCEDSYTOLMNNVPECLSGRSCSSDQVETQACTREQNRICR 84

RESULT 8
US-08-974-022-48
; Sequence 48, Application US/08974022
; Patent No. 6015938
; GENERAL INFORMATION:
; APPLICANT: Boyle, William J.
; APPLICANT: Lacey, David L.
; APPLICANT: Calzone, Frank J.
; APPLICANT: Chang, Ming-Shi
; TITLE OF INVENTION: OSTEOPROTEGERIN
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 Behavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,022
; FILING DATE: 12-DEC-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/577,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-378
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 227 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

; FILING DATE: 08-SEP-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 756-103P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 164 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..164
; OTHER INFORMATION: /note= "TNFR2, see Fig. 5"
; US-08-232-087A-9

Query Match 34.5%; Score 128; DB 3; Length 227;
Best Local Similarity 38.2%; Pred. No. 9.5e-07;
Matches 21; Conservative 5; Mismatches 29; Indels 0; Gaps 0;

QY 6 CRRDSPTCGCPPRHYTQFWNYLRCRYCNVLCGEREEERACHATHNRCRCR 60
Db 67 CTKTSDTVCDSCEDSYTOLMNNVPECLSGRSCSSDQVETQACTREQNRICR 121

RESULT 9
US-08-385-229-2
; Sequence 2, Application US/08385229
; Patent No. 5605690
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Cindy A.
; APPLICANT: Smith, Craig A.
; TITLE OF INVENTION: Method of Treating TNF-Dependent
; TITLE OF INVENTION: Inflammation Using Tumor Necrosis Factor Antagonists
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/385,229
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/946,236
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wight, Christopher L.
; REGISTRATION NUMBER: 31,680
; REFERENCE/DOCKET NUMBER: 2503
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 587-0606
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 461 amino acids
; TYPE: amino acid
; MOLECULE TYPE: protein
; US-08-385-229-2

Query Match 34.5%; Score 128; DB 1; Length 461;
Best Local Similarity 38.2%; Pred. No. 1.9e-06;
Matches 21; Conservative 5; Mismatches 29; Indels 0; Gaps 0;

QY 6 CRRDSPTCGCPPRHYTQFWNYLRCRYCNVLCGEREEERACHATHNRCRCR 60
Db 67 CTKTSDTVCDSCEDSYTOLMNNVPECLSGRSCSSDQVETQACTREQNRICR 121

RESULT 10
US-08-650-000-2
; Sequence 2, Application US/08650000
; Patent No. 5945397
; GENERAL INFORMATION:
```

```

; INFORMATION FOR SEQ ID NO: 1:
;-----
; SEQUENCE CHARACTERISTICS:
;     LENGTH: 486 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;     MOLECULE TYPE: peptide
; US-08-243-010-1

```


5395760-4
LENGTH: 474

Query Match 34.0%; Score 126; DB 5; Length 474;
Best Local Similarity 37.3%; Pred. No. 3.1e-06;
Matches 22; Conservative 5; Mismatches 32; Indels 0; Gaps 0;
Qy 1 FVORPCRDRSPTTCGCPPRHYTOFWNYLERCRYCNVLCGEREEEARACHATHNRACRC 59
Db 63 YVKHFCNKTSDTVCADCEASMYTQWNQFRTCLSCSSCTTDQVEIRACTKQQRVCAC 121

Search completed: January 30, 2001, 16:47:41
Job time: 617 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:45:22 ; Search time 183.67 Seconds
(without alignments)
8.191 Million cell updates/sec

Title: US-09-518-931-2_COPY_132_175

Perfect score: 254

Sequence: 1 CPPGAGVIAPGPSQNTQCQ.....GTFSSSSSQCPHRNCT 44

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq_36: *
- 1: /SIDSL1/gcgdata/geneseq/geneseq/AA1980.DAT:*
 - 2: /SIDSL1/gcgdata/geneseq/geneseq/AA1981.DAT:*
 - 3: /SIDSL1/gcgdata/geneseq/geneseq/AA1982.DAT:*
 - 4: /SIDSL1/gcgdata/geneseq/geneseq/AA1983.DAT:*
 - 5: /SIDSL1/gcgdata/geneseq/geneseq/AA1984.DAT:*
 - 6: /SIDSL1/gcgdata/geneseq/geneseq/AA1985.DAT:*
 - 7: /SIDSL1/gcgdata/geneseq/geneseq/AA1986.DAT:*
 - 8: /SIDSL1/gcgdata/geneseq/geneseq/AA1987.DAT:*
 - 9: /SIDSL1/gcgdata/geneseq/geneseq/AA1988.DAT:*
 - 10: /SIDSL1/gcgdata/geneseq/geneseq/AA1989.DAT:*
 - 11: /SIDSL1/gcgdata/geneseq/geneseq/AA1990.DAT:*
 - 12: /SIDSL1/gcgdata/geneseq/geneseq/AA1991.DAT:*
 - 13: /SIDSL1/gcgdata/geneseq/geneseq/AA1992.DAT:*
 - 14: /SIDSL1/gcgdata/geneseq/geneseq/AA1993.DAT:*
 - 15: /SIDSL1/gcgdata/geneseq/geneseq/AA1994.DAT:*
 - 16: /SIDSL1/gcgdata/geneseq/geneseq/AA1995.DAT:*
 - 17: /SIDSL1/gcgdata/geneseq/geneseq/AA1996.DAT:*
 - 18: /SIDSL1/gcgdata/geneseq/geneseq/AA1997.DAT:*
 - 19: /SIDSL1/gcgdata/geneseq/geneseq/AA1998.DAT:*
 - 20: /SIDSL1/gcgdata/geneseq/geneseq/AA1999.DAT:*
 - 21: /SIDSL1/gcgdata/geneseq/geneseq/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	254	100.0	153	20	Y22222 Human TNFR superfa
2	254	100.0	215	20	W93585 Human hAPO6 protei
3	254	100.0	245	20	Y28449 A human tumour nec
4	254	100.0	271	20	Y42184 Human mFLINT #1 pr
5	254	100.0	273	20	Y42185 Human mFLINT #2 pr
6	254	100.0	300	19	W66102 Amino acid sequenc
7	254	100.0	300	19	W63622 Human tumour necro
8	254	100.0	300	20	Y03099 Human lung TNF-rec
9	254	100.0	300	20	Y42182 Human FLINT #1 pro
10	254	100.0	300	20	Y17479 Mammalian tumour n
11	254	100.0	300	20	Y06817 Human DcR3 polypep
12	254	100.0	300	20	W97749 Human tumour necro

13	254	100.0	300	20	W95082 Orphan receptor (H
14	254	100.0	300	21	Y77458 Human TNF receptor
15	254	100.0	302	20	Y42183 Human FLINT #2 pro
16	142	55.9	401	18	W38343 Rat osteoprotegeri
17	142	55.9	401	18	W38344 Mouse osteoprotege
18	140	55.1	187	17	R99950 Mutated OCIF, OCIF
19	140	55.1	187	21	Y77464 Mutated OCIF, OCIF
20	140	55.1	197	17	R99945 Mutated OCIF, OCIF
21	140	55.1	272	17	R99944 Mutated OCIF, OCIF
22	140	55.1	321	17	R99949 Mutated OCIF, OCIF
23	140	55.1	326	17	R99940 Mutated OCIF, OCIF
24	140	55.1	327	17	R99941 Mutated OCIF, OCIF
25	140	55.1	349	20	W83928 Human FTHMA-070 pa
26	140	55.1	351	17	R99943 Mutated OCIF, OCIF
27	140	55.1	359	17	R99937 Mutated OCIF, OCIF
28	140	55.1	360	17	R99936 Mutated OCIF, OCIF
29	140	55.1	380	17	R99924 Mature osteoclasto
30	140	55.1	390	17	R99357 Human tumour necro
31	140	55.1	391	19	W53238 Human OCIF genome
32	140	55.1	393	17	R99948 Mutated OCIF, OCIF
33	140	55.1	395	19	W57636 Modified Trl recep
34	140	55.1	399	17	R99942 Mutated OCIF, OCIF
35	140	55.1	401	17	R99932 Mutated OCIF, OCIF
36	140	55.1	401	17	R99933 Mutated OCIF, OCIF
37	140	55.1	401	17	R99934 Mutated OCIF, OCIF
38	140	55.1	401	17	R99935 Mutated OCIF, OCIF
39	140	55.1	401	17	R99925 Full length osteoc
40	140	55.1	401	17	R99931 Mutated OCIF, OCIF
41	140	55.1	401	18	W38345 Human osteoprotege
42	140	55.1	401	19	W57635 Trl receptor prote
43	140	55.1	401	19	W53239 Human OCIF genome
44	140	55.1	401	20	Y05742 Tumour necrosis fa
45	140	55.1	401	20	W83926 Human FTHMA-070 pr

ALIGNMENTS

RESULT 1	
Y22222	ID Y22222 standard; Protein; 153 AA.
XX	AC Y22222;
XX	DT 16-SEP-1999 (first entry)
XX	DE Human TNFR superfamily soluble receptor protein sequence.
XX	KW TNFL1; human; TNFR superfamily; tumour necrosis factor ligand; TNF;
XX	KW tumour necrosis factor receptor; TNFR superfamily; cell proliferation;
XX	KW cell differentiation; cytokine production; immunoglobulin; hyperplasia;
XX	KW apoptosis inducer; activated T cell; autoimmune disease; inhibitor;
XX	KW myasthenia gravis; insulin-dependent diabetes mellitus; endotoxin shock;
XX	KW rheumatoid arthritis; multiple sclerosis; systemic lupus erythematosus;
XX	KW tumour; proliferative disorder; neoplasia; dysplasia; immunocompetence;
XX	KW lymphoid organogenesis; bacterial resistance; contact hypersensitivity;
XX	KW delayed type sensitivity; therapy.
OS	Homo sapiens.
XX	PN WO9933980-A2.
XX	PD 08-JUL-1999.
XX	PF 22-DEC-1998; 98WO-US27474.
XX	PR 16-DEC-1998; 98US-0212270.
XX	PR 30-DEC-1997; 97US-0068959.
XX	PA (CHIR) CHIRON CORP.
XX	PI Kassam A, Lamson G, Pot D, Tribouley C;
XX	

DR WPI; 1999-405508/34.
 DR N-PSDB; X84621.
 XX
 PT New tumour necrosis factor ligands, useful for induction of cell
 PT death and/or proliferation of cells
 XX
 XX Claim 1: Page 61; 69pp; English.
 XX
 CC This sequence represents a tumour necrosis factor receptor (TNFR)
 CC superfamily soluble protein of the invention. The invention also relates
 CC to tumour necrosis factor (TNF) ligand (TNFL) family proteins. The TNFL
 CC proteins play regulatory roles in cell proliferation and/or
 CC differentiation, e.g. they can induce production of cytokines,
 CC immunoglobulins, etc. A variety of diseases can be treated by modulating
 CC the activity of TNFL proteins, e.g. they can induce apoptosis of
 CC activated T cells but rescue resting T cell from apoptosis. TNFL
 CC polypeptides can therefore be used to treat autoimmune diseases, such as
 CC myasthenia gravis, insulin-dependent diabetes mellitus, rheumatoid
 CC arthritis, multiple sclerosis, and systemic lupus erythematosus. TNFL
 CC proteins also have tumour stimulating properties, so tumours can be
 CC treated by inhibiting the expression or activity of TNFL. Other
 CC proliferative disorders, such as neoplasias, dysplasias, and hyperplasia
 CC can also be treated using TNFL inhibitors. The TNFL polypeptides and
 CC polynucleotides can also be used to enhance or decrease TNF activity,
 CC thus providing therapeutic benefits such as induction of cell death,
 CC lymphoid organogenesis, or host bacterial resistance, and inhibition of
 CC endotoxic shock, contact hypersensitivity, delayed type sensitivity or
 CC immunocompetence of a transplant recipient. TNF and its receptors play a
 CC major role in host defence and immunosurveillance. As such, there is a
 CC need to identify new members of TNFR families. This invention provides
 CC this need.
 XX
 XX Sequence 153 AA;
 SQ

Query Match 100.0%; Score 254; DB 20; Length 153;
 Best Local Similarity 100.0%; Pred. No. 1.5e-19;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CPPGAGVIAPGTPSNTQCPCPGTFSASSSSSQCPHRNCT 44
 |||||
 Db 48 CPPGAGVIAPGTPSNTQCPCPGTFSASSSSSQCPHRNCT 91

RESULT 2
 W93585
 ID W93585 standard; Protein; 215 AA.
 AC W93585;
 XX
 XX 18-JUN-1999 (first entry)
 DT
 XX
 XX Human hAPO6 protein.
 DE
 XX
 XX Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
 KW developmental abnormality; gestational abnormality; prostate cancer;
 KW APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
 KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
 KW apoptosis; human.
 XX
 OS Homo sapiens.
 XX
 XX WO9911791-A2.
 PN
 XX
 XX 11-MAR-1999.
 PD
 XX
 XX 04-SEP-1998; 98WO-US18393.
 PF
 XX
 XX 05-SEP-1997; 97US-0924634.
 PR
 XX
 XX (UNIW) UNIV WASHINGTON.
 PA
 XX
 XX Chaudhary PM;
 PI

XX WPI; 1999-205191/17.
 DR N-PSDB; X23419.
 XX
 PT New Tumour Necrosis factor family receptor polypeptides and ligands -
 PT useful for diagnosis and treatment of prostate cancer and
 PT developmental or gestational abnormalities
 XX
 XX Claim 29; Fig 9; 156pp; English.
 XX
 CC This invention describes isolated Tumour Necrosis Factor (TNF) family
 CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
 CC their active fragments. APO4 is useful for diagnosing prostate cancer
 CC by determining levels of APO4 in an individual. Prostate cancer can also
 CC be treated using APO4 selective binding agents linked to a therapeutic
 CC moiety. APO4 polypeptides are also useful for identifying selective
 CC binding agents, useful in diagnosis/treatment of disease by binding of
 CC agents to the polypeptide/active fragment which is extracellular, or
 CC expressed on the cell surface. The binding is preferably performed in
 CC vivo. APO4 polypeptides/ active fragments are also useful for screening
 CC for agonists and antagonists by binding and observing the change in APO4
 CC activity. Effective pharmacological agents useful in diagnosis or
 CC treatment of disease are also identified using APO4 polypeptides/active
 CC fragments and APO4 signal transducer molecules that specifically interact
 CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4
 CC activity. The method is performed in vivo or in vitro. APO polypeptides
 CC are all useful as immunogens for preparing antibodies. APO4 is also
 CC useful for diagnosis/treatment of developmental or gestational
 CC abnormalities. APO8 was transfected to human breast carcinoma cell line
 CC MCF-7, and induced apoptosis!
 XX
 XX Sequence 215 AA;
 SQ

Query Match 100.0%; Score 254; DB 20; Length 215;
 Best Local Similarity 100.0%; Pred. No. 2.1e-19;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CPPGAGVIAPGTPSNTQCPCPGTFSASSSSSQCPHRNCT 44
 |||||
 Db 47 CPPGAGVIAPGTPSNTQCPCPGTFSASSSSSQCPHRNCT 90

RESULT 3
 Y28449
 ID Y28449 standard; Protein; 245 AA.
 XX
 XX Y28449;
 XX
 XX 29-SEP-1999 (first entry)
 DT
 XX
 XX A human tumour necrosis factor-R2-like proteins (TR2P)-1.
 DE
 XX
 KW Human tumour necrosis factor-R2-like protein; TR2P; achondroplasia;
 KW osteoporosis; developmental disorder; Cushing's syndrome;
 KW muscular dystrophy; epilepsy; hereditary neuropathy;
 KW Charcot-Marie-Tooth disease; neurofibromatosis; hypothyroidism;
 KW hydrocephalus; seizure disorder; cerebral palsy; spinal bifida;
 KW congenital glaucoma; cataract; sensorineural hearing loss;
 KW reproductive disorder; infertility; ovulatory defect; endometriosis;
 KW autoimmune disorder; ectopic pregnancy; teratogenesis; spermatogenesis;
 KW immunological disorder; AIDS; Addison's disease; allergy; bronchitis;
 KW atherosclerosis; diabetes mellitus; Chron's disease; lupus;
 KW irritable bowel syndrome; multiple sclerosis; infection;
 KW neoplastic disorder; adenocarcinoma; leukaemia; lymphoma; melanoma;
 KW myeloma; sarcoma.
 XX
 OS Homo sapiens.
 XX
 XX WO9931128-A2.
 PN
 XX
 XX 24-JUN-1999.
 PD

XX PF 02-DEC-1998; 98WO-US25649.
XX PR 16-DEC-1997; 97US-0991945.
XX PA (INCY-) INCYTE PHARM INC.
XX PI Au-Young J, Bandman O, Hillman JL, Kaser MR, Tang YT;
XX WPI: 1999-457916/38.
XX DR N-PSDB; X89503.
XX PT New tumour necrosis factor-R2-like protein - useful in the treatment
XX of osteogenesis, developmental, reproductive, immunological and
XX neoplastic disorders
XX PS Claim 1: Fig 1A-C; 81pp; English.
XX
XX The present sequence represents a human tumour necrosis factor-R2-like
XX protein ("R2P"). The protein is used to treat and prevent osteogenesis,
XX developmental, reproductive, immunological and neoplastic disorders, and
XX also to diagnose disorders associated with TR2 protein expression. Such
XX disorders include osteogenesis disorders such as achondroplasia and
XX osteoporosis, developmental disorders such as Cushing's syndrome,
XX muscular dystrophy, and epilepsy, hereditary neuropathies such as
XX Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism,
XX hydrocephalus, seizure disorders such as cerebral palsy and spinal
XX bifida, congenital glaucoma, cataract, or sensorineural hearing loss,
XX reproductive disorders such as infertility, ovulatory defects and
XX endometriosis, autoimmune disorders, ectopic pregnancy and teratogenesis,
XX disruption of spermatogenesis, immunological disorders such as AIDS,
XX Addison's disease, allergies, bronchitis, atherosclerosis, diabetes
XX mellitus, Chron's disease, lupus and irritable bowel syndrome, multiple
XX sclerosis, viral, fungal, helminthic, parasitic and protozoal infections,
XX and neoplastic disorders including adenocarcinoma, leukaemia, lymphoma,
XX melanoma, myeloma, sarcoma, and teratocarcinoma.
XX SQ Sequence 245 AA;

Query Match 100.0%; Score 254; DB 20; Length 245;
Best Local Similarity 100.0%; Pred. No. 2.3e-19;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CPPGAGVIAPGTPSONTCQCPGCTFSASSSSSQCPHRNCT 44
Db 132 cppgagviapgtpsntcqpccppgtfsassssseqcphrnc 175
|||||
RESULT 4
Y42184
ID Y42184 standard; Protein; 271 AA.
XX AC Y42184;
XX DT 17-DEC-1999 (first entry)
XX DE Human mFLINT #1 protein sequence.
XX
XX Human; FLINT; mFLINT; OP3; tumour necrosis factor receptor; FasL;
KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
KW reperfusion-associated injury; aplastic anaemia; differentiation;
KW growth; myelodysplastic syndrome; pancytopenic condition;
KW myocardial ischaemia.
XX
XX Homo sapiens.
XX
XX WO950413-A2.
XX
XX PD 07-OCT-1999.
XX
XX 30-MAR-1999; 99WO-US06797.

XX PR 30-MAR-1998; 98US-0079856.
XX PR 20-MAY-1998; 98US-0086074.
XX PR 09-SEP-1998; 98US-0099643.
XX PR 17-DEC-1998; 98US-0112577.
XX PR 18-DEC-1998; 98US-0112703.
XX PR 18-DEC-1998; 98US-0112933.
XX PR 22-DEC-1998; 98US-0113407.
XX PA (ELIL) LILLY & CO ELI.
XX
XX Bumol TE, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
XX Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
XX Song HY, Wang J, Wu X, Zuckerman SH;
XX WPI: 1999-591319/50.
XX DR N-PSDB; Z25377.
XX
XX Use of mature FLINT for treating acute liver failure, inflammation,
XX cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
XX and proinflammatory activity
XX PS Claim 31; Fig 3; 99pp; English.
XX
XX The present invention describes therapeutic applications of mature FLINT
XX (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
XX (mFLINT), which is a member of the tumour necrosis factor receptor
XX superfamily, is used for treating acute liver failure, inflammation of
XX the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
XX with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
XX injury or disorder such as hypercoagulation (including use with
XX thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
XX or disorder, Type I diabetes, cancer, cell damage or damage to an
XX innocent bystander tissue that is induced by a chemotherapeutic agent or
XX therapeutic irradiation, treating haematopoietic progenitor cells that
XX have been exposed to therapeutic radiation or chemotherapy, aplastic
XX anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
XX also used for promoting the growth or differentiation of a haematopoietic
XX progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
XX resulting from abnormal myocardial ischaemia. The present sequence
XX represents human mFLINT.
XX SQ Sequence 271 AA;

Query Match 100.0%; Score 254; DB 20; Length 271;
Best Local Similarity 100.0%; Pred. No. 2.6e-19;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CPPGAGVIAPGTPSONTCQCPGCTFSASSSSSQCPHRNCT 44
Db 103 cppgagviapgtpsntcqpccppgtfsassssseqcphrnc 146
|||||
RESULT 5
Y42185
ID Y42185 standard; Protein; 273 AA.
XX AC Y42185;
XX DT 17-DEC-1999 (first entry)
XX DE Human mFLINT #2 protein sequence.
XX
XX Human; FLINT; mFLINT; OP3; tumour necrosis factor receptor; FasL;
KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
KW reperfusion-associated injury; aplastic anaemia; differentiation;
KW growth; myelodysplastic syndrome; pancytopenic condition;
KW myocardial ischaemia.
XX
XX Homo sapiens.
XX

PN WO9950413-A2.
 XX
 XX
 PD 07-OCT-1999.
 XX
 XX 30-MAR-1999; 99WO-US06797.
 PF
 XX 30-MAR-1998; 98US-0079856.
 PR 20-MAY-1998; 98US-0086074.
 PR 09-SEP-1998; 98US-0099643.
 PR 17-DEC-1998; 98US-0112577.
 PR 18-DEC-1998; 98US-0112703.
 PR 18-DEC-1998; 98US-0112933.
 PR 22-DEC-1998; 98US-0113407.
 XX
 XX (ELIL) LILLY & CO ELI.
 XX
 XX Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
 PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
 PI Song HY, Wang J, Wu X, Zuckerman SH;
 XX
 XX WPI: 1999-591319/50.
 DR N-PSDB; Z25378.
 XX
 XX Use of mature FLINT for treating acute liver failure, inflammation,
 PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
 PT and proinflammatory activity
 PT
 XX Example 2; Fig 4; 99pp; English.
 PS
 XX The present invention describes therapeutic applications of mature FLINT
 CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
 CC (mFLINT), which is a member of the tumour necrosis factor receptor
 CC superfamily, is used for treating acute liver failure, inflammation of
 CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
 CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
 CC injury or disorder such as hypercoagulation (including use with
 CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
 CC or disorder, Type I diabetes, cancer, cell damage or damage to an
 CC innocent bystander tissue that is induced by a chemotherapeutic agent or
 CC therapeutic irradiation, treating haematopoietic progenitor cells that
 CC have been exposed to therapeutic radiation or chemotherapy, aplastic
 CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
 CC also used for promoting the growth or differentiation of a haematopoietic
 CC progenitor cell or CP34+ cell and preventing damage to a cardiac myocyte
 CC resulting from abnormal myocardial ischaemia. The present sequence
 CC represents human mFLINT.
 XX
 XX Sequence 273 AA;
 SQ
 Query Match 100.0%; Score 254; DB 20; Length 273;
 Best Local Similarity 100.0%; Pred. No. 2.6e-19;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CPBGAGVIAGTFSQNTQCPCPPGTFSSSSSSSQCPHRNCT 44
 Db 105 cpbgagviagtpsgntqcpcppgtfsssssssqcphrnc 148
 |||||||||||||||||||||||||||||||||||||||||||
 RESULT 6
 W66102
 ID W66102 standard; Protein; 300 AA.
 XX
 XX W66102;
 AC
 XX 02-DEC-1998 (first entry)
 DT
 XX Amino acid sequence of tumour necrosis related receptor (TR4).
 DE
 XX Human; tumour necrosis related receptor; TR4; agonist; antagonist;
 KW inhibition; chronic; acute; inflammation; arthritis; septicaemia;
 KW autoimmune disease; transplant rejection; stroke; cancer;
 KW Alzheimer's disease.

XX Homo sapiens.
 OS
 XX EP861850-A1.
 PN
 XX 02-SEP-1998.
 PD
 XX 20-JAN-1998; 98EP-0300382.
 PF
 XX 04-FEB-1997; 97US-0794796.
 PR
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 PA
 XX Emery J, Tan KB, Trunch A, Young PR;
 PI
 XX WPI: 1998-508248/44.
 DR N-PSDB; V07634.
 XX
 XX New DNA encoding tumour necrosis related receptor - used to treat
 PT and prevent e.g. inflammation, arthritis, septicaemia, autoimmune
 PT diseases, transplant rejection, infection, stroke, ischaemia, ARDS,
 PT restenosis, AIDS, bone disorders and cancer
 PT
 XX Claim 1; Fig 1; 21pp; English.
 PS
 XX This is the amino acid sequence of the human tumour necrosis related
 CC receptor (TR4), used in the method of the invention. The TR4 protein
 CC or its agonist can be used to treat a subject in need of enhanced
 CC TR4 polypeptide activity. The antagonist is used to inhibit TR4
 CC polypeptide activity. The active agents can be used for the
 CC treatment and prevention of diseases such as chronic and acute
 CC inflammation, arthritis, septicaemia, autoimmune diseases, transplant
 CC rejection, stroke, cancer, Alzheimer's disease.
 XX
 XX Sequence 300 AA;
 SQ
 Query Match 100.0%; Score 254; DB 19; Length 300;
 Best Local Similarity 100.0%; Pred. No. 2.9e-19;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CPBGAGVIAGTFSQNTQCPCPPGTFSSSSSSSQCPHRNCT 44
 Db 132 cpbgagviagtpsgntqcpcppgtfsssssqcphrnc 175
 |||||||||||||||||||||||||||||||||||||||||||
 RESULT 7
 W63622
 ID W63622 standard; Protein; 300 AA.
 XX
 XX W63622;
 AC
 XX 26-OCT-1998 (first entry)
 DT
 XX Human tumour necrosis factor receptor-6 alpha protein.
 DE
 XX Human tumour necrosis factor receptor-6 alpha; TNFR-6 alpha; TNFR-6 beta;
 KW endothelial cells; keratinocytes; normal prostate; apoptosis;
 KW prostate tumour tissue.
 XX
 XX Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 FH Peptide 1..30
 FT Protein 31..300
 FT /note= "TNFR-6 alpha"
 FT Region 31..282
 FT /note= "Soluble extracellular domain"
 XX
 XX WO9830694-A2.
 PN
 XX 16-JUL-1998.
 PD
 XX

PF 13-JAN-1998; 98WO-US00153.
 XX
 PR 14-JAN-1997; 97US-0035496.
 XX
 FA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ebner R, Feng P, Gentz RL, Ni J, Ruben SM, Yu G;
 XX
 DR WPI: 1998-395142/34.
 XX
 DR N-PSDB; V39085.
 XX
 PT Human tumour necrosis factor receptors 6-alpha and 6-beta - used in
 XX the diagnosis of immune system-related disorder(s)
 XX
 PS Claim 20; Fig 1; 91pp; English.
 XX
 CC The present sequence represents the human tumour necrosis factor
 CC receptor-6 alpha (TNFR-6 alpha) protein. The invention also provides
 CC for the TNFR-6 beta protein (W63623). TNFR-6 alpha and TNFR-6 beta are
 CC members of the tumour necrosis factor receptor (TNFR) family. TNFRs
 CC are expressed in endothelial cells, keratinocytes, normal prostate and
 CC prostate tumour tissue. For a number of disorders of these cells,
 CC particularly of the immune system, substantially altered (whether
 CC increased or decreased) levels of TNFR-6 alpha and/or TNFR-6 beta gene
 CC expression can be detected, therefore the TNFR-6 alpha and TNFR-6 beta
 CC polypeptides, nucleic acids and antibodies are claimed to be useful in
 CC the diagnosis of such disorders. Mutations of the TNFR-6 alpha and
 CC TNFR-6 beta genes can also be detected. The TNFR polypeptides are
 CC also claimed to be useful for identifying ligands which may be useful
 CC in the treatment of apoptosis related disorders.
 XX
 SQ Sequence 300 AA;

Query Match 100.0%; Score 254; DB 19; Length 300;
 Best Local Similarity 100.0%; Pred. No. 2.9e-19;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CPPGAGVIAPGTPSQNTQCPCPGTFSASSSSSQCPHRNCT 44
 Db 132 cppdgaviaptsgntqtcpcpgtfsassssseqcphrnc 175
 |||||
 RESULT 8
 Y03099
 ID Y03099 standard; Protein; 300 AA.
 AC Y03099;
 XX
 XX 09-DEC-1999 (first entry)
 DT Human lung TNF-receptor protein.
 DE
 XX
 KW Tumour necrosis factor; TNF; TNF receptor; human; lung; gene therapy;
 KW detection; immunoassay; diagnosis; disease; immune system; tumour;
 KW osteogenic system; cardiovascular system; central nervous system; asthma;
 KW peripheral nervous systems; transplant incompatibility; antitumor;
 KW rheumatoid arthritis; antiasthmatic; antiarthritic.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT CDS 134..1036
 FT /*tag= a
 FT /product= "TNF-receptor"
 XX
 XX DE19809978-A1.
 XX
 PD 16-SEP-1999.
 XX
 XX 09-MAR-1998; 98DE-1009978.
 XX
 XX 09-MAR-1998; 98DE-1009978.
 PR

XX (BADI) BASF AG.
 XX
 XX Kroeger B;
 PI
 XX WPI: 1999-519473/44.
 XX
 DR N-PSDB; Z09998.
 XX
 PT New soluble member of tumor necrosis factor receptor family, useful for
 XX identification specific modulators and for treating disease e.g. tumors
 PT
 XX
 XX Claim 1; Page 8-9; 10pp; German.
 PS
 XX
 CC This invention describes a novel tumour necrosis factor (TNF) receptor
 CC (I) isolated from human lung tissue. (I) is used: (i) to raise specific
 CC antibodies (Ab); (ii) to screen for specific (antagonists or ligands
 CC (A), potential therapeutic agents; and (iii) therapeutically (optionally
 CC expressed from a gene therapy vector) in conditions associated with a
 CC deficit of (I). Ab are used: (a) for qualitative or quantitative
 CC detection of (I) in standard immunoassays (for diagnosis of disease, or
 CC susceptibility, or for monitoring); and (b) as therapeutic inhibitors in
 CC cases where (I) is overexpressed. Nucleic acid (II) that encodes (I) is
 CC used: (A) for recombinant production of (I); (B) also its oligonucleotide
 CC fragments, in standard hybridization and/or amplification assays; (C) as
 CC source of antisense molecules or ribozymes; and (D) to produce transgenic
 CC animals (for studying (patho)physiology of (I)). Diseases possibly
 CC associated with under- or over-expression of (I) are those of the immune,
 CC osteogenic, cardiovascular and central or peripheral nervous systems,
 CC tumors, transplant incompatibility, asthma and rheumatoid arthritis. The
 CC products of the invention have antitumor, antiasthmatic and
 CC antiarthritic activity. This sequence represents the TNF-receptor of the
 CC invention.
 XX
 SQ Sequence 300 AA;

Query Match 100.0%; Score 254; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 2.9e-19;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CPPGAGVIAPGTPSQNTQCPCPGTFSASSSSSQCPHRNCT 44
 Db 132 cppgagviaptsgntqtcpcpgtfsassssseqcphrnc 175
 |||||
 RESULT 9
 Y42182
 ID Y42182 standard; Protein; 300 AA.
 AC Y42182;
 XX
 XX 17-DEC-1999 (first entry)
 DT Human FLINT #1 protein sequence.
 DE
 XX
 KW Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
 KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
 KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
 KW reperfusion-associated injury; aplastic anaemia; differentiation;
 KW growth; myelodysplastic syndrome; pancytopenic condition;
 KW myocardial ischaemia.
 XX
 OS Homo sapiens.
 XX
 XX W09950413-A2.
 XX
 XX 07-OCT-1999.
 PD
 XX
 XX 30-MAR-1999; 99WO-US06797.
 PF
 XX 30-MAR-1998; 98US-0079856.
 PR
 XX 20-MAY-1998; 98US-0086074.
 PR

```

PR 09-SEP-1998; 98US-0099643.
PR 17-DEC-1998; 98US-0112577.
PR 18-DEC-1998; 98US-0112703.
PR 18-DEC-1998; 98US-0112933.
PR 22-DEC-1998; 98US-0113407.
XX
XX (ELIL ) LILLY & CO ELI.
XX
XX Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
PI Hui KY, Khartonenkov A, Mizrahi J, Na S, Noblitt TW, Reldy CA;
PI Song HY, Wang J, Wu X, Zuckerman SH;
XX
XX WPI; 1999-591319/50.
DR N-PSDB; Z25375.
XX
XX use of mature FLINT for treating acute liver failure, inflammation,
PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
PT and proinflammatory activity
XX
XX Claim 30; Fig 1; 99pp; English.
XX
XX The present invention describes therapeutic applications of mature FLINT
CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
CC (mFLINT), which is a member of the tumour necrosis factor receptor
CC superfamily, is used for treating acute liver failure, inflammation of
CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
CC injury or disorder such as hypercoagulation (including use with
CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
CC or disorder, type I diabetes, cancer, cell damage or damage to an
CC innocent bystander tissue that is induced by a chemotherapeutic agent or
CC therapeutic irradiation, treating haematopoietic progenitor cells that
CC have been exposed to therapeutic radiation or chemotherapy, aplastic
CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
CC also used for promoting the growth or differentiation of a haematopoietic
CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
CC resulting from abnormal myocardial ischaemia. The present sequence
CC represents human FLINT.
XX
XX Sequence 300 AA;
SQ
Query Match 100.0%; Score 254; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.9e-19;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CPPGAGVIAPGTPSQNTQCPCPPGTFSSSSSSSQCPHNRNCT 44
Db 132 cpggagviapgtspntqtcpcppgtfsssssssqcqhnrnct 175
|||||
RESULT 10
Y17479
ID Y17479 standard; Protein; 300 AA.
XX
XX Y17479;
AC
XX
XX 02-AUG-1999 (first entry)
DT
XX
XX Mammalian tumour necrosis factor receptor OPG-2.
DE
XX
XX Tumour necrosis factor receptor; TNF receptor; OPG-2; Paget's disease;
KW osteopenic disorder; osteoclast activity; primary osteoporosis;
KW hyperglycaemia; osteolytic metastasis; immune response; cancer.
XX
XX Mammalia.
OS
XX
XX WO9926977-A1.
PN
XX
XX 03-JUN-1999.
PD
XX
XX 24-NOV-1998; 98WO-US25065.
PF
XX

```

```

PR 17-FEB-1998; 98US-0074896.
PR 24-NOV-1997; 97US-0066446.
XX
XX (BIOJ ) BIOGEN INC.
XX
XX Tschopp J;
PI
XX
XX WPI; 1999-347693/29.
DR N-PSDB; X76052.
XX
XX New tumour necrosis factor family receptor OPG-2
PT
XX
XX Claim 1; Page 18; 22pp; English.
XX
XX The present sequence represents a mammalian tumour necrosis factor
CC receptor, designated OPG-2. OPG-2, is a member of the tumour necrosis
CC factor receptor family, and can be used: (i) to raise specific
CC antibodies (Ab), (ii) to treat osteopenic disorders associated with
CC excessive osteoclast activity, e.g. primary osteoporosis, Paget's
CC disease, hyperglycaemia of malignancy, or osteolytic metastases; (iii)
CC for affinity purification of cognate ligands, and (iv) to screen for
CC ligands (antagonists or agonists). Ab, or other OPG-2 blocking agents
CC such as soluble forms of the protein, are used to prevent, or reduce
CC severity of, an immune response, and for treating cancer. They can also
CC be used in diagnostic assays. The nucleic acid sequence encoding OPG-2
CC can be used as a probe to isolate related sequences from other species.
XX
XX Sequence 300 AA;
SQ
Query Match 100.0%; Score 254; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.9e-19;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CPPGAGVIAPGTPSQNTQCPCPPGTFSSSSSQCPHNRNCT 44
Db 132 cpggagviapgtspntqtcpcppgtfsssssssqcqhnrnct 175
|||||
RESULT 11
Y06817
ID Y06817 standard; Protein; 300 AA.
XX
XX Y06817;
AC
XX
XX 24-JUN-1999 (first entry)
DT
XX
XX Human DcR3 polypeptide.
DE
XX
XX DcR3 polypeptide; tumour necrosis factor receptor; TNFR; Fas ligand;
KW apoptosis; T cell mediated immune response; allergy; asthma; cancer;
KW rheumatoid arthritis; Crohn's disease; guest vs. host disease; human;
KW gene therapy.
XX
XX Homo sapiens.
OS
XX
XX WO9914330-A1.
PN
XX
XX 25-MAR-1999.
PD
XX
XX 18-SEP-1998; 98WO-US19661.
PF
XX
XX 30-JUL-1998; 98US-0094640.
PR
XX
XX 18-SEP-1997; 97US-0059288.
PR
XX
XX (GETH ) GENENTECH INC.
PA
XX
XX Ashkenazi AJ, Botstein D, Dodge KH, Goddard A, Gurney AL;
PI Kim KJ, Lawrence DA, Pitti R, Roy MA, Tumas DB;
PI Wood WI;
XX
XX WPI; 1999-244032/20.
DR N-PSDB; X32744.

```


XX Dcr3 polypeptide related to tumor necrosis factor receptor
PT Claim 5; Fig 1; 88pp; English.
PS
XX This represents a human Dcr3 polypeptide, a homologue of tumour necrosis
CC factor receptor (TNFR) polypeptide. Host cells containing a vector
CC comprising the Dcr3 nucleic acid can be used for the recombinant
CC expression of the protein. Dcr3 binds to Fas ligand, so it (or its
CC chimeras) are useful for modulating apoptosis in mammalian cells, also
CC other Fas-ligand induced activities, particularly to inhibit T cell
CC mediated immune responses, e.g. in treatment of allergy, asthma,
CC rheumatoid arthritis, Crohn's disease, guest vs. host disease etc. Dcr3
CC may also be used to identify specific binding proteins, potential
CC inhibitors. Antibodies against Dcr3 are used to treat cancer,
CC specifically of the lung and colon, also in diagnosis and for affinity
CC purification of the protein. Detecting mutations in the gene for Dcr3 is
CC also used to diagnose cancer, or predisposition to it. Dcr3 nucleic acid
CC is useful as hybridization probe to detect genomic or related sequences;
CC for chromosome and gene mapping; as source of antisense sequences; for
CC expression of recombinant Dcr3 and to generate transgenic animals (for
CC development and screening of therapeutic agents), also for in vivo or
CC ex vivo gene therapy.
XX
SQ Sequence 300 AA;

Query Match 100.0%; Score 254; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.9e-19;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CPPGAGVIAPGPSQNTQCPCPGTFSASSSSSQCPHRNCT 44
|||||
Db 132 cpggagviapgtpsqntcpcpgptfsassssseqcphrnc 175

RESULT 12
W97749
ID W97749 standard; Protein; 300 AA.
XX
AC W97749;
XX
DT 21-MAY-1999 (first entry)
XX
DE Human tumour necrosis factor receptor ZTNFR-5.
XX
KW ZTNFR-5; tumour necrosis factor receptor; TNFR; human;
KW cell maturation; bone cell regulation.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..23
FT /note= "signal peptide"
FT Protein 24..300
FT /note= "mature protein"
FT Domain 24..194
FT /note= "extracellular domain"
FT Region 49..71
FT /note= "cysteine-rich pseudo-repeat 1"
FT Region 72..113
FT /note= "cysteine-rich pseudo-repeat 1"
FT Region 114..151
FT /note= "cysteine-rich pseudo-repeat 1"
FT Region 152..194
FT /note= "cysteine-rich pseudo-repeat 1"
XX
XX WO9904001-A1.
XX
PD 28-JAN-1999.
XX
XX 21-JUL-1998; 98WO-US15072.
XX

PR 21-JUL-1997; 97US-0053203.
XX (ZYMO) ZYMOGENETICS INC.
PA
XX Farrah TM;
PI
XX WPI; 1999-132245/11.
DR N-PSDB; X07226.
XX
PT Novel tumour necrosis factor receptor ZTNFR5 - useful for
PT regulating maturation of TNF-ligand bearing cells
XX
XX Claim 1; Page 84-85; 109pp; English.
XX
CC This polypeptide comprises a new, secreted tumour necrosis factor
CC receptor (see W97749), designated ZTNFR-5. Novel ZTNFR-5 encoding
CC polynucleotides and polypeptides were initially identified by
CC querying an expressed sequence tag (EST) database for sequences
CC homologous to conserved motifs within the TNF receptor family.
CC Based on this search, a contig of 16 ESTs (see X07226) was
CC constructed. ZTNFR-5 polypeptides comprise 4 cysteine-rich repeats
CC (see also W97750-55) that are homologous to other TNF receptors, in
CC particular the soluble, secreted TNF receptor osteoprotegerin.
CC ZTNFR-5 polypeptide can be prepared by recombinant methods. The
CC polypeptide, especially the extracellular domain, can be used to
CC generate a soluble variant of ZTNFR-5. The polypeptides and
CC nucleic acids can be used to screen for ligands, agonists and
CC antagonists of ZTNFR-5. The polypeptides can be used in bone cell
CC regulation and to regulate the maturation of TNF ligand-bearing
CC cells such as T- or B-cells, lymphocytes, peripheral blood
CC mononuclear cells, polymorphonuclear leukocytes, fibroblasts or
CC haematopoietic cells.
XX
SQ Sequence 300 AA;

Query Match 100.0%; Score 254; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.9e-19;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CPPGAGVIAPGPSQNTQCPCPGTFSASSSSSQCPHRNCT 44
|||||
Db 132 cpggagviapgtpsqntcpcpgptfsassssseqcphrnc 175

RESULT 13
W95082
ID W95082 standard; Protein; 300 AA.
XX
AC W95082;
XX
DT 20-MAY-1999 (first entry)
XX
DE Orphan receptor (HUMAN NTR-1) polypeptide.
XX
KW HUMAN NTR-1; orphan receptor; osteoprotegerin; OPG; TNFR; human;
KW tumour necrosis factor receptor; muscle disorder; bone mass; screening;
KW muscle metabolism; binding agent; cognate ligand.
XX
OS Homo sapiens.
XX
FN WO9907738-A2.
XX
PD 18-FEB-1999.
XX
XX 04-AUG-1998; 98WO-US16202.
XX
XX 06-AUG-1997; 97US-0054869.
XX
XX (PROC) PROCTER & GAMBLE CO.
XX (REGE-) REGENERON PHARM INC.
XX
PI Masiakowski PJ, Morris J, Valenzuela DM;

XX WPI; 1999-167365/14.
 DR N-PSDB; X22300.
 XX
 XX Novel orphan human receptor polypeptide and nucleic acid - useful as
 PT diagnostic reagents and for treatment of muscle disorders
 XX
 XX
 PS Claim 7; Page 21; 23pp; English.
 XX
 XX This represents a HUMAN NTR-1 polypeptide, a novel orphan receptor. The
 CC protein is related to osteoprotegerin (OPG) and to tumour necrosis factor
 CC receptor (TNFR). Host cells transformed with a vector comprising the
 CC HUMAN NTR-1 nucleic acid are used for the recombinant expression of the
 CC protein. HUMAN NTR-1 proteins and antibodies immuno specific for the
 CC protein are useful for diagnosis and treatment of humans and animals,
 CC especially muscle disorders, as the receptor is involved in regulation of
 CC bone mass and muscle metabolism. HUMAN NTR-1 receptors are also useful
 CC for screening for novel binding agents, and cognate ligands, which may be
 CC used to treat disorders associated with HUMAN NTR-1 imbalance.
 XX
 XX Sequence 300 AA;
 SQ
 Query Match 100.0%; Score 254; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 2.9e-19;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CPPGAGVIAPGTPSQNTQCPCPGTFFSASSSSSQCPHRNCT 44
 DB 132 CPPGAGVIAPGTPSQNTQCPCPGTFFSASSSSSQCPHRNCT 175
 RESULT 14
 Y77458
 ID Y77458 standard; Protein: 300 AA.
 AC Y77458;
 XX
 XX 05-JUN-2000 (first entry)
 DT
 XX Human TNF receptor-like protein, HDTEA84.
 DE
 XX
 XX TNF receptor family; tumour necrosis factor; HDTEA84; HSLJD37R;
 KW Rank-like protein; RANKL; immune disorder; inflammation; allergy;
 KW immunosuppressant; antiarthritic; antirheumatoid; antiinflammatory;
 KW dermatological; antithyroid.
 XX
 XX Homo sapiens.
 OS
 XX
 XX WO200001817-A2.
 PN
 XX
 XX 13-JAN-2000.
 PD
 XX
 XX 06-JUL-1999; 99WO-US12366.
 PF
 XX
 XX 06-JUL-1998; 98US-0110938.
 PR
 XX 13-JUL-1998; 98US-0114466.
 PR
 XX 23-JUL-1998; 98US-0093897.
 PR
 XX 12-AUG-1998; 98US-0132968.
 PR
 XX 18-AUG-1998; 98US-0136214.
 PR
 XX 11-SEP-1998; 98US-0099999.
 PR
 XX
 XX (SCHE) SCHERING CORP.
 PA
 XX
 XX Bates EEM, Lebecque SJE, Murphy EE, Mattison JD, Gormah DM;
 PI Hedrick JA, Wang L, Zlotnik A, Murgolo NJ, Greene JR, Johnston JA;
 PI Bazan JF, Mahony D, Lees EM;
 XX
 XX WPI; 2000-171015/15.
 DR N-PSDB; Z92404.
 DR
 XX
 XX New isolated mammalian genes, used to develop products for treating
 PT e.g. immune, inflammatory or allergic abnormalities, cancers or

PT degenerative conditions -
 XX
 XX Claim 24; Page 157; 218pp; English.
 XX
 CC The invention relates to a number of primate and/or rodent proteins, and
 CC the genes which encode them. The invention encompasses human dendritic
 CC cell prostaglandin transporter (DC-PGT); the TNF (tumour necrosis
 CC factor) receptor family-related proteins HDTEA84, HSLJD37R and RANKL;
 CC human CC chemokine HCC5; human deubiquitinating proteins Dub1 and Dub
 CC 12; human MD-1 and human and murine MD-2 proteins which exhibit the
 CC properties of ligands for proteins comprising a leucine-rich motif (LRR);
 CC human cyclin E2; cDNAs encoding these proteins; and antibodies against
 CC these proteins. The proteins can be used for modulating the physiology or
 CC development of a cell. They can be used to mediate uptake of substrates
 CC (e.g., prostaglandin-like molecules), to modulate or mediate cellular
 CC interactions (e.g., induce or prevent trafficking, proliferation, or
 CC differentiation of cells), or are intracellular proteins which are
 CC important in various cellular processes such as the deubiquitination of
 CC proteins or cell cycle regulation. The products can be used for
 CC treating medical conditions such as immune, inflammatory or allergic
 CC disorders, or abnormal cellular proliferation, for example, cancers or
 CC degenerative conditions. They can be used to modulate immune responses in
 CC disease states e.g., autoimmune disorders, including rheumatoid
 CC arthritis, systemic lupus erythematosus, Hashimoto's autoimmune
 CC thyroiditis, as well as acute and chronic inflammatory responses in which
 CC T cell activation, expansion, and/or immunological T cell memory play an
 CC important role. Sequences Y77458-Y77461 and Y77465-Y77468 represent
 CC TNF receptor family-related proteins. Y77458 is the human protein
 CC HDTEA84, Y77459-Y77461 are human HSLJD37R proteins, Y77465 is
 CC murine Rank-like protein RANKL, and Y77466-Y77468 are human RANKL
 CC proteins.
 XX
 XX Sequence 300 AA;
 SQ
 Query Match 100.0%; Score 254; DB 21; Length 300;
 Best Local Similarity 100.0%; Pred. No. 2.9e-19;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CPPGAGVIAPGTPSQNTQCPCPGTFFSASSSSSQCPHRNCT 44
 DB 132 CPPGAGVIAPGTPSQNTQCPCPGTFFSASSSSSQCPHRNCT 175
 RESULT 15
 Y42183
 ID Y42183 standard; Protein: 302 AA.
 AC Y42183;
 XX
 XX 17-DEC-1999 (first entry)
 DT
 XX Human FLINT #2 protein sequence.
 DE
 XX
 XX Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
 KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
 KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
 KW reperfusion-associated injury; aplastic anaemia; differentiation;
 KW growth; myelodysplastic syndrome; pancytopenic condition;
 KW myocardial ischaemia.
 XX
 XX Homo sapiens.
 OS
 XX
 XX WO9950413-A2.
 PN
 XX
 XX 07-OCT-1999.
 PD
 XX
 XX 30-MAR-1999; 99WO-US06797.
 PF
 XX
 XX 30-MAR-1998; 98US-0079856.
 PR
 XX 20-MAY-1998; 98US-0086074.
 PR
 XX 09-SEP-1998; 98US-0099643.
 PR
 XX 17-DEC-1998; 98US-0112577.
 PR

```

PR 18-DEC-1998; 98US-0112703.
PR 18-DEC-1998; 98US-0112933.
PR 22-DEC-1998; 98US-0113407.
XX
PA (ELIL ) LILLY & CO ELI.
XX
PI Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
PI Song HY, Wang J, Wu X, Zuckerman SH;
XX
DR WPT: 1999-591319/50.
DR N-PSDB: 225376.
XX
PT Use of mature FLINT for treating acute liver failure, inflammation,
PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
PT and proinflammatory activity
XX
PS Example 2; Fig 2: 99pp; English.
XX
CC The present invention describes therapeutic applications of mature FLINT
CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
CC (mFLINT), which is a member of the tumour necrosis factor receptor
CC superfamily, is used for treating acute liver failure, inflammation of
CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
CC injury or disorder such as hypercoagulation (including use with
CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
CC or disorder, Type I diabetes, cancer, cell damage or damage to an
CC innocent bystander tissue that is induced by a chemotherapeutic agent or
CC therapeutic irradiation, treating haematopoietic progenitor cells that
CC have been exposed to therapeutic radiation or chemotherapy, aplastic
CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
CC also used for promoting the growth or differentiation of a haematopoietic
CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
CC resulting from abnormal myocardial ischaemia. The present sequence
CC represents human FLINT.
XX
SQ Sequence 302 AA;

```

```

Query Match      100.0%; Score 254; DB 20; Length 302;
Best Local Similarity 100.0%; Pred. No. 2.9e-19;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CPPGAGVIAPGTPSONTCQCPGCTESASSSSSEQCOPHRNCT 44
   |||||||
Db 134 cppgagviapgtcpsqntqcpptgfsassssseqcphrnc 177

```

Search completed: January 30, 2001, 16:45:22
Job time: 562 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run On: January 30, 2001, 16:50:21 ; Search time 149.64 seconds
(without alignments)
19.965 Million cell updates/sec

Title: US-09-518-931-2_COPY_132_175

Perfect score: 254

Sequence: 1 CPPGAGVIADPTSQNTQC.....GTFSASSSSQCPHNRNCT 44

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_66:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	136	53.5	461	1 A35356	tumor necrosis fac
2	126	49.6	459	2 I48854	gene murine tumour
3	126	49.6	474	2 B38634	tumor necrosis fac
4	107	42.1	289	2 A46515	B cell-associated
5	107	42.1	305	2 A46476	B cell activation
6	104	40.9	277	2 A60771	tumor necrosis fac
7	94.5	37.2	435	2 I54182	OX40 homolog hum
8	93	36.6	277	2 I37552	nerve growth facto
9	90.3	35.6	416	1 JN0006	OX40 antigen precu
10	89	35.0	271	2 S12783	nerve growth facto
11	87.5	34.4	425	1 GGHUN	nerve growth facto
12	86.5	34.1	427	1 A26431	gene ox40 protein
13	86	33.9	272	2 I48700	T2 protein - rabbi
14	83.5	32.9	325	2 B43692	CD30 antigen precu
15	83.5	32.9	595	2 A42086	T2 protein - myxom
16	78.5	30.9	326	1 GQVZML	tumor necrosis fac
17	76.5	30.1	461	1 GQRTT1	receptor-type prot
18	76	29.9	1013	2 I50615	hypothetical prote
19	75	29.5	348	2 T28623	G2R protein - vari
20	75	29.5	349	2 D72175	gene 34R protein -
21	75	29.5	349	2 D36858	protein-tyrosine k
22	75	29.5	998	2 S37627	tumor necrosis fac
23	72.5	28.5	454	1 GQMT1	tumor necrosis fac
24	72.5	28.5	454	2 I57826	receptor-like tyro
25	72	28.3	893	2 S51603	receptor tyrosine
26	72	28.3	898	2 S47489	receptor-like tyro
27	72	28.3	981	2 S51604	receptor tyrosine
28	72	28.3	1005	2 S49015	probable membrane
29	71	28.0	797	2 T34966	

RESULT 1

A35356

tumor necrosis factor receptor type 2 precursor - human
N:Alternate names: 75K tumor necrosis factor receptor
C:Species: Homo sapiens (man)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: A35356; A36475; A48416; A36007; A23666; B35010; I38094

R:Smith, C.A.; Davis, T.; Anderson, D.; Solam, L.; Beckmann, M.P.; Jerzy, R.; Dower,

Science 248, 1019-1023, 1990

A:Title: A receptor for tumor necrosis factor defines an unusual family of cellular a

A:Reference number: A35356; MUID:90260639

A:Accession: A35356

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-461 <SM>

A:Cross-references: GB:M33315; NID:q189185; PIDN:AAA59929.1; PID:q189186

R:Kohn, T.; Brewer, M.T.; Baker, S.L.; Schwartz, P.E.; King, M.W.; Hale, K.K.; Squir

Proc. Natl. Acad. Sci. U.S.A. 87, 8331-8335, 1990

A:Title: A second tumor necrosis factor receptor gene product can shed a naturally oc

A:Reference number: A36475; MUID:91045991

A:Accession: A36475

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-195, 'R', 197-461 <KH>

A:Cross-references: GB:M55994; GB:M38549; NID:g339757; PIDN:AAA36755.1; PID:g339758

R:Dembic, Z.; Loetscher, H.; Gubler, U.; Pan, Y.C.; Lahm, H.W.; Gentz, R.; Brockhaus,

Cytokine 2, 231-237, 1990

A:Title: Two human TNF receptors have similar extracellular, but distinct intracellular

A:Reference number: A48416; MUID:91370690

A:Accession: A48416

A>Status: preliminary

A:Molecule type: mRNA; protein

A:Residues: 23-461 <DEM>

A:Cross-references: GB:S63368; NID:q235648; PIDN:AAB19824.1; PID:q235649

A>Note: sequence extracted from NCBI backbone (NCBIN:63368, NCBIP:63371)

R:Heiler, R.A.; Song, K.; Onasch, M.A.; Fischer, W.H.; Chang, D.; Ringold, G.M.

Proc. Natl. Acad. Sci. U.S.A. 87, 6151-6155, 1990

A:Title: Complementary DNA cloning of a receptor for tumor necrosis factor and demons

A:Reference number: A36007; MUID:90349572

A:Accession: A36007

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 116-140, 'P', 142-195, 'R', 197-362, 'T', 364-461 <HEL>

A:Cross-references: GB:M35857; NID:g339751; PIDN:AAA63262.1; PID:g339752

R:Loetscher, H.; Schlaeger, E.J.; Lahm, H.W.; Pan, Y.C.E.; Lesslauer, W.; Brockhaus,

J. Biol. Chem. 265, 20131-20138, 1990

A:Title: Purification and partial amino acid sequence analysis of two distinct tumor

A:Reference number: A23666; MUID:91056048

A:Accession: A23666

A>Status: preliminary

A:Molecule type: protein

A:Residues: 23-40; 65-69; 136-141; 300-306 <LOE>

```

R:Engelmann, H.; Novick, D.; Wallach, D.
J. Biol. Chem. 265, 1531-1536, 1990
A:Title: Two tumor necrosis factor-binding proteins purified from human urine. Evidence
A:Reference number: A35010; MUID:90110215
A:Accession: B35010
A:Molecule type: protein
A:Status: preliminary
A:Residues: 27-31 <ENG>
R:Kuhnert, P.; Kemper, O.; Wallach, D.
Gene 150, 381-386, 1994
A:Title: Cloning, sequencing and partial functional characterization of the 5' region of
A:Reference number: I38094; MUID:95121934
A:Accession: I38094
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-37 <RES>
A:Cross-references: EMBL:X80021; NID:g666044; PIDN:CAA56324.1; PID:g825701
C:Genetics:
A:Gene: GDB:TNFR2
A:Cross-references: GDB:I25914; OMIM:191191
A:Map position: lp36.2-lp36.2
A:Introns: 26/3
A:Note: the list of introns is incomplete
C:Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology
C:Keywords: duplication; glycoprotein; receptor; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-416/Product: tumor necrosis factor receptor type 2 #status experimental <MAT>
F:40-76/Domain: NGF receptor repeat homology <NG1>
F:78-119/Domain: NGF receptor repeat homology <NG2>
F:120-162/Domain: NGF receptor repeat homology <NG3>
F:164-201/Domain: NGF receptor repeat homology <NG4>
F:262-279/Domain: transmembrane #status predicted <TMN>
F:280-461/Domain: intracellular #status predicted <INT>
F:171,193/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 53.5%; Score 136; DB 1; Length 461;
Best Local Similarity 51.2%; Pred. No. 1.5e-05;
Matches 22; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 1 CPPGAGVIAPGTPSQNTQCPCPPGTFSSASSSSQCCPHRNC 43
DB 143 CRPGFGVAPGRTSDVVKCPKCAPGTFTSTSDICRPHQIC 185

RESULT 2
I48854
gene murine tumour necrosis factor receptor 2 protein - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 23-Jul-1999
C:Accession: I48854
R:Powell, E.E.; Wicker, L.S.; Peterson, L.B.; Todd, J.A.
Mamm. Genome 5, 726-727, 1994
A:Title: Allelic variation of the type 2 tumor necrosis factor receptor gene.
A:Reference number: I48854; MUID:95178848
A:Accession: I48854
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-459 <RES>
A:Cross-references: EMBL:X76401; NID:g433830; PIDN:CAA53981.1; PID:g433831
C:Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology
F:151-188/Domain: NGF receptor repeat homology <NGF>

Query Match 49.6%; Score 126; DB 2; Length 459;
Best Local Similarity 47.7%; Pred. No. 1.4e-05;
Matches 21; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 1 CPPGAGVIAPGTPSQNTQCPCPPGTFSSASSSSQCCPHRNC 44
DB 130 CGPGFGVASSRAPNGVNLKACAPGTFSDTSTSDVCRPHRCS 173

```

```

RESULT 3
B38634
tumor necrosis factor receptor type 2 precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C:Accession: B38634; A40254; S54816
R:Lewis, M.; Tartaglia, L.A.; Lee, A.; Bennett, G.L.; Rice, G.C.; Wong, G.H.W.; Chen,
Proc. Natl. Acad. Sci. U.S.A. 88, 2830-2834, 1991
A:Title: Cloning and expression of cDNAs for two distinct murine tumor necrosis facto
A:Reference number: A38634; MUID:91187885
A:Accession: B38634
A:Molecule type: mRNA
A:Residues: 1-474 <LEW>
A:Cross-references: GB:M60469; NID:g199827; PIDN:AAA39752.1; PID:g199828
R:Goodwin, R.G.; Anderson, D.; Jerzy, R.; Davis, T.; Brannan, C.I.; Copeland, N.G.; J
Mol. Cell. Biol. 11, 3020-3026, 1991
A:Title: Molecular cloning and expression of the type 1 and type 2 murine receptors f
A:Reference number: A40254; MUID:91246168
A:Accession: A40254
A:Molecule type: mRNA
A:Residues: 1-474 <GOO>
A:Cross-references: GB:M60469; NID:g199827; PIDN:AAA39752.1; PID:g199828
R:Kissmerghis, M.; Fellows, R.; Feldmann, M.; Chernajovsky, Y.
submitted to the EMBL Data Library, May 1995
A:Description: Characterization of the promoter region of the murine p75-TNF receptor
A:Reference number: S54816
A:Accession: S54816
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-22 <NIS>
A:Cross-references: EMBL:X87128; NID:g809043; PIDN:CAA60618.1; PID:g809044
C:Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology
C:Keywords: cytokine receptor; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-474/Product: tumor necrosis factor receptor type 2 #status predicted <MAT>
F:40-77/Domain: NGF receptor repeat homology <NG1>
F:79-120/Domain: NGF receptor repeat homology <NG2>
F:166-203/Domain: NGF receptor repeat homology <NG4>

Query Match 49.6%; Score 126; DB 2; Length 474;
Best Local Similarity 47.7%; Pred. No. 1.5e-05;
Matches 21; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 1 CPPGAGVIAPGTPSQNTQCPCPPGTFSSASSSSQCCPHRNC 44
DB 145 CGPGFGVASSRAPNGVNLKACAPGTFSDTSTSDVCRPHRCS 188

RESULT 4
A46515
B cell-associated surface molecule CD40, short splice form - mouse
C:Species: Mus musculus (house mouse)
C:Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 26-May-2000
C:Accession: A46515
R:Grimaldi, J.C.; Torres, R.; Kozak, C.A.; Chang, R.; Clark, E.A.; Howard, M.; Cockay
J. Immunol. 149, 3921-3926, 1992
A:Title: Genomic structure and chromosomal mapping of the murine CD40 gene.
A:Reference number: A46515; MUID:93094586
A:Accession: A46515
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-289 <GRI>
A:Cross-references: GB:M83312; NID:g1553058; PIDN:AAB08705.1; PID:g1553059; GB:M94126
A:Experimental source: BALB/c, liver
A:Note: sequence extracted from NCBI backbone (NCBIP:120357)
C:Comment: For an alternative splice form, see PIR:A46476
C:Superfamily: CD27 antigen; NGF receptor repeat homology
C:Keywords: alternative splicing
F:105-144/Domain: NGF receptor repeat homology <NGF>

Query Match 42.1%; Score 107; DB 2; Length 289;

```

Best Local Similarity 46.5%; Pred. NO. 0.00074;
Matches 20; Conservative 5; Mismatches 18; Indels 0; Gaps 0;

QY 1 CPPGAGVIAPTPSQNTQCPCPGTFSASSSSSEQCQPHRNC 43
DB 125 CIPGFGVMEMATETDTVCHPCPGVGFESNQSLFEKCYPTWTC 167

RESULT 5
A46476
B cell-associated surface molecule CD40, long splice form - mouse
C:Species: Mus musculus (house mouse)
C:Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 26-May-2000
C:Accession: A46476
R:Torres, R.M.; Clark, E.A.
J. Immunol. 148, 620-626, 1992

A:Title: Differential increase of an alternatively polyadenylated mRNA species of murine A:Reference number: A46476; MUID:92105763
A:Accession: A46476
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-305 <TOR>
A:Cross-references: GB:M3312; NID:g1553058
A:Note: sequence extracted from NCBI backbone (NCBIN:75206, NCBIP:75207)
A:Comment: this translation is not annotated in GenBank entry MUSCD40A, release 113.0
C:Superfamily: CD27 antigen; NGF receptor repeat homology
C:Keywords: alternative splicing; transmembrane protein
F:105-144/Domain: NGF receptor repeat homology <NGF>

Query Match 42.1%; Score 107; DB 2; Length 305;
Best Local Similarity 46.5%; Pred. NO. 0.00077;
Matches 20; Conservative 5; Mismatches 18; Indels 0; Gaps 0;

QY 1 CPPGAGVIAPTPSQNTQCPCPGTFSASSSSSEQCQPHRNC 43
DB 125 CIPGFGVMEMATETDTVCHPCPGVGFESNQSLFEKCYPTWTC 167

RESULT 6
A60771
B-cell activation protein CD40 precursor - human
N:Alternate names: B-cell surface antigen Bp50
C:Species: Homo sapiens (man)
C:Date: 03-Jun-1993 #sequence_revision 03-Feb-1994 #text_change 21-Jul-2000
C:Accession: S04460; A60771
R:Stamenkovic, I.; Clark, E.A.; Seed, B.
EMBO J. 8, 1403-1410, 1989
A:Title: A B-lymphocyte activation molecule related to the nerve growth factor receptor
A:Reference number: S04460; MUID:89356608
A:Accession: S04460
A:Molecule type: mRNA
A:Residues: 1-277 <STA>
A:Cross-references: EMBL:X60592; NID:g29850; PIDN:CAA43045.1; PID:g29851
R:Braesch-Andersen, S.; Paulie, S.; Koho, H.; Nika, H.; Aspenstroem, P.; Perlmann, P.
J. Immunol. 142, 562-567, 1989
A:Title: Biochemical characteristics and partial amino acid sequence of the receptor-like A:Reference number: A60771; MUID:89093941
A:Accession: A60771
A:Molecule type: protein
A:Residues: 21-50 <BRA>
A:Experimental source: Burkitt lymphoma cell line Raji
C:Genetics:
A:Gene: GDB:CD40
A:Cross-references: GDB:215268; OMIM:109535
A:Map position: 20q12-20q13.2
C:Superfamily: CD27 antigen; NGF receptor repeat homology
C:Keywords: B-cell; glycoprotein; phosphoprotein; surface antigen; transmembrane protein
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-277/Product: B-cell activation protein CD40 #status experimental <MAT>
F:21-193/Domain: extracellular #status predicted <EXT>
F:194-215/Domain: transmembrane #status predicted <TMW>

F:216-277/Domain: intracellular #status predicted <CYT>
F:153,180/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 40.9%; Score 104; DB 2; Length 277;
Best Local Similarity 46.5%; Pred. NO. 0.0014;
Matches 20; Conservative 5; Mismatches 18; Indels 0; Gaps 0;

QY 1 CPPGAGVIAPTPSQNTQCPCPGTFSASSSSSEQCQPHRNC 43
DB 125 CSPGFGVKQIATGVSDTICEPCPGVGFESNVSAFCKHPWTSC 167

RESULT 7
154182
tumor necrosis factor receptor 2-related protein - human
C:Species: Homo sapiens (man)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 17-Mar-2000
C:Accession: I54182
R:Baens, M.; Chaffanet, M.; Cassiman, J.J.; Van den Berghe, H.; Marynen, P.
Genomics 16, 214-218, 1993
A:Title: Construction and evaluation of a hncDNA library of human 12p transcribed seq A:Reference number: I54182; MUID:93252381
A:Accession: I54182
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-435 <RES>
A:Cross-references: GB:L04270; NID:g339761; PIDN:AAA36757.1; PID:g339762
C:Genetics:
A:Gene: GDB:LTBR
A:Cross-references: GDB:1230195; OMIM:600979
A:Map position: 12p13.3-12p13.1
C:Superfamily: tumor necrosis factor receptor type 1; NGF receptor repeat homology

Query Match 37.2%; Score 94.5; DB 2; Length 435;
Best Local Similarity 40.9%; Pred. NO. 0.018;
Matches 18; Conservative 3; Mismatches 22; Indels 1; Gaps 1;

QY 1 CPPGA-GVIAPTPSQNTQCPCPGTFSASSSSSEQCQPHRNC 43
DB 148 CPPGTEAEKDEVGKGNHCVCKAGHFQNTSPSARCQPHTRC 191

RESULT 8
I37552
OX40 homolog - human
C:Species: Homo sapiens (man)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 11-Jan-2000
C:Accession: I37552
R:Latzka, U.; Durkop, H.; Schnittger, S.; Ringeling, J.; Eitelbach, F.; Hummel, M.; Fo Eur. J. Immunol. 24, 677-683, 1994
A:Title: The human OX40 homolog: cDNA structure, expression and chromosomal assignment A:Reference number: I37552; MUID:94170844
A:Accession: I37552
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-277 <RES>
A:Cross-references: EMBL:X75962; NID:g472957; PIDN:CAA33576.1; PID:g472958
C:Superfamily: CD27 antigen; NGF receptor repeat homology

Query Match 36.6%; Score 93; DB 2; Length 277;
Best Local Similarity 45.5%; Pred. NO. 0.017;
Matches 20; Conservative 7; Mismatches 13; Indels 4; Gaps 2;

QY 1 CPPGAGVIAPTPSQNTQCPCPGTFSASSSSSEQCQPHRNC 44
DB 46 CRPGNGMWSRCSRSQNTVCRCPGPG-FYNDVVSSKPKRP---CT 85

RESULT 9
JN0006

QY 1 CPPCAGVIA---PQT-PSQNT-----QCPCPPGTFTSASSSSSEQCPHNRCT 44
| | : | | | | | | | | | | | | | | | | | | | |
DB 94 CTPTEDIVCCRPCTQPRQDSSHKLGVDCVPCPGHF--SPGSNQAOKPWTNCT 145

RESULT 11
GQHUN
nerve growth factor receptor precursor, low affinity - human
N;Alternate names: NGF receptor
C;Species: Homo sapiens (man)
C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 22-Jun-1999
C;Accession: A25218; A60204; S21689; I57638
R;Johnson, D.; Lananhan, A.; Buck, C.R.; Sehgal, A.; Morgan, C.; Mercer, E.; Bothwell,
Cell 47, 545-554, 1986
A;Title: Expression and structure of the human NGF receptor.
A;Reference number: A25218; MUID:87051725
A;Accession: A25218
A:Molecule type: mRNA
A;Residues: 1-427 <JOH>
A;Cross-references: GB:M14764; NID:g189204; PIDN:AAB59544.1; PID:g189205
R;Marano, N.; Dietzschold, B.; Earley Jr., J.J.; Schattelman, G.; Thompson, S.; Grob,
J. Neurochem. 48, 225-232, 1987
A;Title: Purification and amino terminal sequencing of human melanoma nerve growth fa
A;Reference number: A60204; MUID:87085574
A;Accession: A60204
A:Molecule type: protein
A;Residues: 29-31,'T','33'-42,'TT','45-46,'TX','50-51,'XX','54-56 <MAR>
A;Experimental source: melanoma cell line A875
A;Note: This sequence has been corrected by a note added in proof to follow the nucle
R;Visavajjhala, P.; Leszyk, J.D.; Lin-Goeke, J.; Ross, A.H.
Arch. Biochem. Biophys. 294, 244-252, 1992
A;Title: Structural domains of the extracellular domain of human nerve growth factor
A;Reference number: S21689; MUID:92198017
A;Accession: S21689
A>Status: preliminary
A:Molecule type: protein
A;Residues: 183-208 <VIS>
R;Sehgal, A.; Patil, N.; Chao, M.
Mol. Cell. Biol. 8, 3160-3167, 1988
A;Title: A constitutive promoter directs expression of the nerve growth factor recept
A;Reference number: I57638; MUID:89096903
A;Accession: I57638
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A;Residues: 1-22 <RES>
A;Cross-references: GB:M21621; NID:g189206; PIDN:AAA36363.1; PID:g189207
C;Comment: This receptor is found on sensory and sympathetic neurons, on neuroblastom
C;Comment: The cysteine-rich region of the extracellular domain may form part or all
C;Comment: This protein is thought to form a high-affinity receptor when it associate
C;Comment: This receptor undergoes both N- and O-linked glycosylation

Query Match 34.4%; Score 87.5; DB 1; Length 427;
Best Local Similarity 36.4%; Pred. No. 0.088;
Matches 16; Conservative 8; Mismatches 17; Indels 3; Gaps 1;

A;Cross-references: GDB:131547; OMIM:153243
A;Map position: lp36-lp36
C;Superfamily: NGF receptor repeat homology
C;Keywords: glycoprotein; growth factor receptor; transmembrane protein
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-383/Domain: extracellular #status predicted <EXT>
F;384-407/Domain: transmembrane #status predicted <TM>
F;408-595/Domain: intracellular #status predicted <CYT>
F;101,276/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 32.9%; Score 83.5; DB 2; Length 595;
Best Local Similarity 44.7%; Pred. No. 0.29;
Matches 17; Conservative 5; Mismatches 15; Indels 1; Gaps 1;

QY 1 CPPGAGVIAPGTPSONTCQCPPTGTFSSSSSEQC 38
||| | ||| :|| | | | | : | | |
Db 131 CPAGMIVKFFGTAKNTVCEPASPGV-SPACASPENCK 167

Search completed: January 30, 2001, 16:50:21
Job time: 711 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 17:05:56 ; Search time 80.04 Seconds
(without alignments)
17.753 Million cell updates/sec

Title: US-09-518-931-2_COPY_132_175

Perfect score: 254

Sequence: 1 CPPGAGVIAGTPSQNTQCO.....GTFSSSSSQEQPHRNCT 44

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	136	53.5	461	1	TNR2_HUMAN
2	126	49.6	474	1	TNR2_MOUSE
3	107	42.1	289	1	CD40_MOUSE
4	104	40.9	277	1	CD40_HUMAN
5	94.5	37.2	435	1	TNRC_HUMAN
6	93	36.6	277	1	OX40_MOUSE
7	91.5	36.0	415	1	TNRC_HUMAN
8	91	35.8	269	1	CD40_BOVIN
9	90.5	35.6	416	1	NGFR_CHICK
10	89	35.0	271	1	OX40_RAT
11	87.5	34.4	427	1	NGFR_HUMAN
12	86.5	34.1	425	1	NGFR_RAT
13	86	33.9	272	1	OX40_MOUSE
14	83.5	32.9	325	1	VTR2_SFVKA
15	83.5	32.9	595	1	CD30_HUMAN
16	78.5	30.9	326	1	VTR2_MXXVL
17	76.5	30.1	461	1	TNR1_RAT
18	76	29.9	1013	1	EPAS5_CHICK
19	75	29.5	349	1	VC22_VARV
20	75	29.5	998	1	EPB3_HUMAN
21	72.5	28.5	454	1	TNR1_MOUSE
22	72	28.3	1005	1	EPAS5_RAT
23	69	27.2	981	1	EPAS3_BRARE
24	68	26.8	1037	1	EPAS5_HUMAN
25	68	26.8	2768	1	THYG_MOUSE
26	67.5	26.6	255	1	41BB_MOUSE
27	67.5	26.6	455	1	TNR1_HUMAN
28	67.5	26.6	1475	1	N153_HUMAN
29	67	26.4	103	1	VA53_VACCC
30	67	26.4	103	1	VA53_VACCV
31	67	26.4	984	1	EPB1_CHICK
32	67	26.4	993	1	EPB3_MOUSE
33	67	26.4	1468	1	N153_RAT

34	66.5	26.2	471	1	TNR1_BOVIN	019131 bos taurus
35	66	26.0	301	1	CC02_CAEL	P17656 caenorhabdi
36	65.5	25.8	122	1	VC22_VACCC	P21106 vaccinia vi
37	65.5	25.8	3164	1	TEGU_HSV11	P10220 herpes simp
38	65	25.6	250	1	CD27_MOUSE	P41272 mus musculu
39	65	25.6	461	1	TNR1_PIG	P50555 sus scrofa
40	65	25.6	755	1	COMP_RAT	P35444 rattus norv
41	65	25.6	987	1	EPB2_COTJA	Q90344 coturnix co
42	64	25.2	984	1	EPB1_HUMAN	P54762 homo sapien
43	64	25.2	2481	1	UN52_CAEL	Q05561 caenorhabdi
44	63.5	25.0	417	1	WSL1_HUMAN	Q93038 h wsl-1 pro
45	63.5	25.0	1416	1	YN81_CAEL	Q03610 caenorhabdi

ALIGNMENTS

RESULT	1
TNR2_HUMAN	
ID	TNR2_HUMAN STANDARD; PRT; 461 AA.
AC	P20333;
DT	01-FEB-1991 (Rel. 17, Created)
DT	01-AUG-1991 (Rel. 19, Last sequence update)
DT	01-OCT-2000 (Rel. 40, Last annotation update)
DE	TUMOR NECROSIS FACTOR RECEPTOR 2 PRECURSOR (TUMOR NECROSIS FACTOR BINDING PROTEIN 2) (TBPII) (P80) (TNF-R2) (P75) (CD120B) (ETANERCEPT).
DE	TNRSF1B OR TNFR2 OR TNFR.
GN	Homo sapiens (Human).
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=90260639; PubMed=2160731;
RA	Smith C.A., Davis T., Anderson D., Solam L., Beckmann M.P., Jerzy R., Dower S.K., Cosman D., Goodwin R.G.;
RA	"A receptor for tumor necrosis factor defines an unusual family of cellular and viral proteins.";
RT	Science 248:1019-1023(1990).
RL	[2]
RN	SEQUENCE FROM N.A.
RP	MEDLINE=91045991; PubMed=2172983;
RX	Kohn T., Brewer M.T., Baker S.L., Schwartz P.E., King M.W., Hale K.K., Squires C.H., Thompson R.C., Vannice J.L.;
RA	"A second tumor necrosis factor receptor gene product can shed a naturally occurring tumor necrosis factor inhibitor.";
RT	Proc. Natl. Acad. Sci. U.S.A. 87:8331-8335(1990).
RL	[3]
RN	SEQUENCE FROM N.A.
RP	MEDLINE=96299745; PubMed=8661109;
RX	Beltlinger C.P., White P.S., Maris J.M., Sulman E.P., Jensen S.J., Lepaslier D., Stallard B.J., Goeddel D.V., Desauvage F.J., Brodeur G.M.;
RA	"Physical mapping and genomic structure of the human TNFR2 gene.";
RT	Genomics 35:94-100(1996).
RL	[4]
RN	SEQUENCE OF 116-461 FROM N.A., AND PARTIAL SEQUENCE.
RP	MEDLINE=90349572; PubMed=2166946;
RX	Heller R.A., Song K., Onasch M.A., Fischer W.H., Chang D., Ringold G.M.;
RA	"Complementary DNA cloning of a receptor for tumor necrosis factor and demonstration of a shed form of the receptor.";
RT	Proc. Natl. Acad. Sci. U.S.A. 87:6151-6155(1990).
RL	[5]
RN	SEQUENCE OF 27-31.
RP	MEDLINE=90110215; PubMed=2153136;
RX	Engelmann H., Novick D., Wallach D.;
RA	"Two tumor necrosis factor-binding proteins purified from human urine. Evidence for immunological cross-reactivity with cell surface tumor necrosis factor receptors.";
RT	J. Biol. Chem. 265:1531-1536(1990).
RL	[6]
RN	SEQUENCE OF 22-40; 65-69; 136-141; 300-306 AND 346-362.
RP	MEDLINE=91056048; PubMed=2173696;
RX	

RA Loetscher H., Schlaeager E.J., Lahm H.-W., Pan Y.-C.E., Lesslauer W.,
 RA Brockhaus M.;
 RT "Purification and partial amino acid sequence analysis of two
 RT distinct tumor necrosis factor receptors from HL60 cells.";
 J. Biol. Chem. 265:20131-20138(1990).
 RN [7]
 RN CHARACTERIZATION.
 RP MEDLINE-93016040; PubMed-1328224;
 RA Pennica D., Lam V.T., Mize N.K., Weber R.F., Lewis M., Fendly B.M.,
 RA Lipari M.T., Goeddel D.V.;
 RA "Biochemical properties of the 75-kDa tumor necrosis factor receptor.
 RT Characterization of ligand binding, internalization, and receptor
 RT phosphorylation.";
 RL J. Biol. Chem. 267:21172-21178(1992).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 419-428 IN TRAF2 COMPLEX.
 RX MEDLINE-99221490; PubMed-10206649;
 RA Park Y.C., Burkitt V., Villa A.R., Tong L., Wu H.;
 RT "Structural basis for self-association and receptor recognition of
 RT human TRAF2.";
 RL Nature 398:533-538(1999).
 CC -I- FUNCTION: RECEPTOR FOR TNF-ALPHA. HIGH AFFINITY FOR TNA-ALPHA AND
 CC APPROXIMATELY 5-FOLD LOWER AFFINITY FOR TNF-BETA.
 CC -I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -I- PTM: PHOSPHORYLATED; MAINLY ON SERINE RESIDUES WITH A VERY LOW
 CC LEVEL ON THREONINE RESIDUES.
 CC -I- PHARMACEUTICAL: AVAILABLE UNDER THE NAME ENBREL (IMMUNEX AND
 CC WYETH-AYERST). USED TO TREAT MODERATE TO SEVERE RHEUMATOID
 CC ARTHRITIS (RA). ENBREL CONSIST OF THE EXTRACELLULAR LIGAND-BINDING
 CC PORTION OF TNFR2 LINKED TO AN IMMUGLOBULIN FC CHAIN. IT BINDS TO
 CC TNF-ALPHA AND CONTAINS ITS INTERACTIONS WITH RECEPTORS.
 CC -I- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 CC -I- DATABASE: NAME=PROW; NOTE=CD guide CD120b entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd120b.htm".
 CC -I- DATABASE: NAME=Enbrel; NOTE=Clinical information on Enbrel;
 CC WWW="http://www.enbrelinfo.com/".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M32315; AAA59929.1; -;
 DR EMBL; M35857; AAA63262.1; -;
 DR EMBL; U52165; AAC50622.1; -;
 DR EMBL; U52156; AAC50622.1; JOINED.
 DR EMBL; U52157; AAC50622.1; JOINED.
 DR EMBL; U52158; AAC50622.1; JOINED.
 DR EMBL; U52159; AAC50622.1; JOINED.
 DR EMBL; U52160; AAC50622.1; JOINED.
 DR EMBL; U52161; AAC50622.1; JOINED.
 DR EMBL; U52162; AAC50622.1; JOINED.
 DR EMBL; U52163; AAC50622.1; JOINED.
 DR EMBL; U52164; AAC50622.1; JOINED.
 DR EMBL; M55994; AAA36755.1; -;
 DR PIR; A35356; A35356.
 DR PIR; A36007; A36007.
 DR PIR; A36475; A36475.
 DR PIR; B35010; B35010.
 DR PIR; A23666; A23666.
 DR PDB; 1CA9; 12-APR-99.
 DR MIM; 191191; -;
 DR INTERPRO; IPR001368; -;
 DR PRAM; PF00020; TNFR_c6; 4.
 DR PROSITE; PS00652; TNFR_NGFR_1; 2.
 DR PROSITE; PS0050; TNFR_NGFR_2; 4.
 KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal;
 KW Phosphorylation; Pharmaceutical; 3D-structure.
 FT SIGNAL 1 22
 FT CHAIN 23 461 TUMOR NECROSIS FACTOR RECEPTOR 2.

FT DOMAIN 23 257 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 287 POTENTIAL.
 FT DOMAIN 288 461 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 39 201 4 X TNFR-CYS.
 FT REPEAT 39 76 TNFR-CYS 1.
 FT REPEAT 77 118 TNFR-CYS 2.
 FT REPEAT 119 162 TNFR-CYS 3.
 FT REPEAT 163 201 TNFR-CYS 4.
 FT DISULFID 40 53 BY SIMILARITY.
 FT DISULFID 54 67 BY SIMILARITY.
 FT DISULFID 57 75 BY SIMILARITY.
 FT DISULFID 78 93 BY SIMILARITY.
 FT DISULFID 96 110 BY SIMILARITY.
 FT DISULFID 100 118 BY SIMILARITY.
 FT DISULFID 120 126 BY SIMILARITY.
 FT DISULFID 134 143 BY SIMILARITY.
 FT DISULFID 137 161 BY SIMILARITY.
 FT DISULFID 164 179 BY SIMILARITY.
 FT CARBOHYD 171 171 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 193 193 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 141 141 R -> P (IN REF. 4).
 FT CONFLICT 196 196 R -> M (IN REF. 1 AND 3).
 FT CONFLICT 363 363 A -> T (IN REF. 4).
 SQ SEQUENCE 461 AA; 48316 MW; 603B580ECD67636F CRC64;
 Query Match 53.5%; Score 136; DB 1; Length 461;
 Best Local Similarity 51.2%; Pred. No. 6.2e-07;
 Matches 22; Conservative 9; Mismatches 12; Indels 0; Gaps 0;
 QY 1 CPFGAGVIAPGTSTQCPGPGTFSASSSSSSSQCPHRNC 43
 Db 143 CRPGFGVARGTSTQCPGPGTFSASSSSSSSQCPHRNC 185
 RESULT 2
 TNFR2_MOUSE STANDARD; PRT; 474 AA.
 ID TNFR2_MOUSE AC P25119; P97893;
 AC P25119; P97893;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE TUMOR NECROSIS FACTOR RECEPTOR 2 PRECURSOR (TNF-R2) (P75).
 GN TNFRSF1B OR TNFR2 OR TNFR-2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-91187885; PubMed-1849278;
 RA Lewis M., Tartaglia L.A., Lee A., Bennett G.L., Rice G.C.,
 RA Wong G.H., Chen E.Y., Goeddel D.V.;
 RT "Cloning and expression of cDNAs for two distinct murine tumor
 RT necrosis factor receptors demonstrate one receptor is species
 RT specific.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:2830-2834(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-91246168; PubMed-1645445;
 RA Goodwin R.G., Anderson D., Jerzy R., Brannan C.I.,
 RA Copeland N.G., Jenkins N.A., Smith C.A.;
 RT "Molecular cloning and expression of the type 1 and type 2 murine
 RT receptors for tumor necrosis factor.";
 RL Mol. Cell. Biol. 11:3020-3026(1991).
 RN [3]
 RP SEQUENCE OF 1-26 FROM N.A.
 RC STRAIN=NOD;
 RA Jacob C.O., Liu J.;
 RN Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
 [4]
 RN SEQUENCE OF 1-22 FROM N.A.
 RC TISSUE=LIVER;
 RA Kissonegrhis M., Fellows R., Feldmann M., Chernaiovsky Y.;

```
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: RECEPTOR FOR TNF-ALPHA.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M60469; AAA39752.1; -.
DR EMBL; M59378; AAA40463.1; -.
DR EMBL; U39488; AAA85021.1; -.
DR EMBL; X87128; CAA60618.1; -.
DR PIR; B38634; B38634.
DR HSSP; P19438; INCF.
DR MGD; MGI:1314883; TNFRSF1B.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00500; TNFR_NGFR_2; 3.
KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 22
FT CHAIN 23 474 TUMOR NECROSIS FACTOR RECEPTOR 2.
FT DOMAIN 23 258 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 259 288 POTENTIAL.
FT DOMAIN 289 474 CYTOPLASMIC (POTENTIAL).
FT REPEAT 39 203 4 X TNFR-CYS.
FT REPEAT 39 77 TNFR-CYS 1.
FT REPEAT 78 119 TNFR-CYS 2.
FT REPEAT 120 164 TNFR-CYS 3.
FT REPEAT 165 203 TNFR-CYS 4.
FT DISULFID 40 54 BY SIMILARITY.
FT DISULFID 55 68 BY SIMILARITY.
FT DISULFID 58 76 BY SIMILARITY.
FT DISULFID 79 94 BY SIMILARITY.
FT DISULFID 97 111 BY SIMILARITY.
FT DISULFID 101 119 BY SIMILARITY.
FT DISULFID 121 127 BY SIMILARITY.
FT DISULFID 136 145 BY SIMILARITY.
FT DISULFID 139 163 BY SIMILARITY.
FT DISULFID 166 181 BY SIMILARITY.
FT CARBOHYD 69 69 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 474 AA; 50319 MW; 462EAE398C4D6563 CRC64;

Query Match 49.6%; Score 126; DB 1; Length 474;
Best Local Similarity 47.7%; Pred. No. 6.3e-06; Indels 0; Gaps 0;
Matches 21; Conservative 9; Mismatches 14;

OY 1 CPPGAGVIAPGTPSNTQCPCPPGTFSSASSSSSQCPHRNC 44
Db 145 CGPFGVASSRAPNGVLCACAPGTFSDTSSTDVCRPHRCS 188

RESULT 3
CD40_MOUSE STANDARD; PRT; 289 AA.
AC P27512;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CD40L RECEPTOR PRECURSOR (B-CELL SURFACE ANTIGEN CD40) (BP50) (CDW40).
GN TNFRSF5 OR CD40.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
```

```
RX MEDLINE-92105763; PubMed-1370315;
RA Torres R.M., Clark E.A.;
RT Differential increase of an alternatively polyadenylated mRNA
RT species of murine CD40 upon B lymphocyte activation.;
J. Immunol. 148:620-626(1992).
RN [2]
RN REVIEWS.
RC STRAIN=BALB/C;
RA Torres R.M.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=LIVER;
RX MEDLINE-93094586; PubMed-1281194;
RA Grimaldi J.C., Torres R., Kozak C.A., Chang R., Clark E.A.,
RA Howard M., Cockayne D.A.;
RT Genomic structure and chromosomal mapping of the murine CD40 gene.;
J. Immunol. 149:3921-3926(1992).
RL -!- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD40L.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M83312; AAB08705.1; -.
DR EMBL; M94126; AAA37404.1; -.
DR EMBL; M94129; AAA37404.1; JOINED.
DR EMBL; M94128; AAA37404.1; JOINED.
DR EMBL; M94127; AAA37404.1; JOINED.
DR PIR; A46476; A46476.
DR HSSP; P25942; 1CDF.
DR MGD; MGI-88336; TNFRSF5.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 1.
DR PROSITE; PS00500; TNFR_NGFR_2; 4.
KW Receptor; B-cell; Glycoprotein; Transmembrane; Repeat; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 289 CD40L RECEPTOR.
FT DOMAIN 20 193 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 194 215 POTENTIAL.
FT DOMAIN 216 289 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 25 187 4 X TNFR-CYS.
FT REPEAT 25 60 TNFR-CYS 1.
FT REPEAT 61 103 TNFR-CYS 2.
FT REPEAT 104 144 TNFR-CYS 3.
FT REPEAT 145 187 TNFR-CYS 4.
FT CARBOHYD 153 153 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 289 AA; 32111 MW; C791CB6D2FEA574E CRC64;

Query Match 42.1%; Score 107; DB 1; Length 289;
Best Local Similarity 46.5%; Pred. No. 0.00031;
Matches 20; Conservative 5; Mismatches 18; Indels 0; Gaps 0;

OY 1 CPPGAGVIAPGTPSNTQCPCPPGTFSSASSSSSQCPHRNC 43
Db 125 CIPGFGVMEMATETDTVCHPCPGVGFSSNQSLFKRCYPTWSC 167

RESULT 4
CD40_HUMAN STANDARD; PRT; 277 AA.
ID CD40_HUMAN
AC P25942;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
RP SEQUENCE FROM N.A.
```

CD40L RECEPTOR PRECURSOR (B-CELL SURFACE ANTIGEN CD40) (BP50) (CDW40)
 (TNFRSF5 OR CD40).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=89356608; PubMed=2475341;
 RA Stamenkovic I., Clark E.A., Seed B.;
 RT "A B-lymphocyte activation molecule related to the nerve growth
 factor receptor and induced by cytokines in carcinomas";
 RL EMBO J. 8:1403-1410(1989).
 [2]
 RN 3D-STRUCTURE MODELING OF 24-144.
 RX MEDLINE=97189482; PubMed=9037712;
 RA Bajorath J., Aruffo A.;
 RT "Construction and analysis of a detailed three-dimensional model of
 the ligand binding domain of the human B cell receptor CD40";
 RL Proteins 27:59-70(1997).
 [3]
 RN 3D-STRUCTURE MODELING OF 26-186 IN COMPLEX WITH CD40L.
 RX MEDLINE=98266353; PubMed=9605317;
 RA Singh J., Garber E., van Vlijmen H., Karpusas M., Hsu Y.-M.,
 RT "The role of polar interactions in the molecular recognition of CD40L
 with its receptor CD40";
 RL Protein Sci. 7:1124-1135(1998).
 CC -!- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD40L.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -!- TISSUE SPECIFICITY: B-CELLS AND IN PRIMARY CARCINOMAS.
 CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD40 entry;
 WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd40.htm".
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X60592; CAA43045.1; --
 DR PIR: S04460; S04460.
 DR PDB: 1CDF; 01-APR-97.
 DR MIM: 109535; --
 DR INTERPRO: IPR001368; --
 DR PFAM: PF00020; TNFR_C6; 4.
 DR PROSITE: PS00652; TNFR_NGFR_1; 1.
 DR PROSITE: PS50050; TNFR_NGFR_2; 4.
 DR Receptor; B-cell; Glycoprotein; Transmembrane; Repeat; Signal;
 KW 3D-structure. 1
 FT SIGNAL 1 19
 FT CHAIN 20 277
 FT DOMAIN 20 193
 FT TRANSMEM 194 215
 FT DOMAIN 216 277
 FT DOMAIN 25 187
 FT REPEAT 25 60
 FT REPEAT 61 103
 FT REPEAT 104 144
 FT REPEAT 145 187
 FT REPEAT 26 37
 FT DISULFID 38 51
 FT DISULFID 41 59
 FT DISULFID 62 77
 FT DISULFID 83 103
 FT DISULFID 105 119
 FT DISULFID 111 116
 FT DISULFID 125 143
 FT DISULFID 153 153
 FT CARBOHYD 180 180
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT N-LINKED (GLCNAC. .) (POTENTIAL).
 FT
 SQ SEQUENCE 277 AA; 30619 MW; BC8776EC2C4A5680 CRC64;
 Query Match 40.9%; Score 104; DB 1; Length 277;
 Best Local Similarity 46.5%; Pred. No. 0.0006;
 Matches 20; Conservative 5; Mismatches 18; Indels 0; Gaps 0;
 QY 1 CPPGAGVIACTPSONTCOCPCPGTFSASSSSSESCQCPHRNC 43
 DB 125 CSPGFGVKQIATGVSDTICEPCVPVGVFFSNVSSAFKCHPWTS 167
 RESULT 5
 TNRC_HUMAN STANDARD; PRT; 435 AA.
 ID TNRC_HUMAN
 AC P36941;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE LYMPHOTOXIN-BETA RECEPTOR PRECURSOR (TUMOR NECROSIS FACTOR RECEPTOR
 DE 2 RELATED PROTEIN) (TUMOR NECROSIS FACTOR C RECEPTOR).
 GN LTBR OR TNFR OR TNFRSF3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE=93252381; PubMed=8486360;
 RA Baens M., Chaffanet M., Cassiman J.J., den Berghe H., Marynen P.;
 RT "Construction and evaluation of a hcdDNA library of human 12p
 RT transcribed sequences derived from a somatic cell hybrid";
 RL Genomics 16:214-218(1993).
 [2]
 RN FUNCTION.
 RP MEDLINE=94225209; PubMed=8171323;
 RA Crowe P.D., van Arsdele T.L., Walter B.N., Ware C.F., Hession C.,
 RA Ehrenfels B., Browning J.L., Din W.S., Goodwin R.G., Smith C.A.;
 RT "A lymphotoxin-beta-specific receptor";
 RL Science 264:707-710(1994).
 CC -!- FUNCTION: RECEPTOR FOR THE LYMPHOTOXIN-BETA. POSSIBLE FUNCTION IN
 CC IMMUNE DEVELOPMENT.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL: L04270; AAA36757.1; --
 DR HSSP: P25942; 1CDF.
 DR MIM: 600979; --
 DR INTERPRO: IPR001368; --
 DR PFAM: PF00020; TNFR_C6; 4.
 DR PROSITE: PS00652; TNFR_NGFR_1; 2.
 DR PROSITE: PS50050; TNFR_NGFR_2; 3.
 DR Receptor; Transmembrane; Glycoprotein; Repeat; Signal;
 FT SIGNAL 1 30
 FT CHAIN 31 435
 FT DOMAIN 31 227
 FT TRANSMEM 228 248
 FT DOMAIN 249 435
 FT DOMAIN 42 211
 FT REPEAT 42 81
 FT REPEAT 82 124
 FT REPEAT 125 168
 FT REPEAT 169 211
 FT DISULFID 43 58
 FT DISULFID 59 72
 FT BY SIMILARITY.
 FT BY SIMILARITY.

```
FT DISULFID 62 80 BY SIMILARITY.
FT DISULFID 83 98 BY SIMILARITY.
FT DISULFID 101 116 BY SIMILARITY.
FT DISULFID 104 124 BY SIMILARITY.
FT DISULFID 126 132 BY SIMILARITY.
FT DISULFID 139 148 BY SIMILARITY.
FT DISULFID 142 167 BY SIMILARITY.
FT DISULFID 170 185 BY SIMILARITY.
FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 435 AA; 46709 MW; 6246266022F656F CRC64;

Query Match 37.2%; Score 94.5; DB 1; Length 435;
Best Local Similarity 40.9%; Pred. No. 0.0079;
Matches 18; Conservative 3; Mismatches 22; Indels 1; Gaps 1;

QY 1 CPPGA-GVIAPGTPSQNTQCPCPPGTFSSSSSSSQCCQPHRNC 43
||||| : ||||| : ||||| : ||||| : ||||| :
Db 148 CPPGTEAEKDEVGKNNHCVPCKAGHFQNTSSPSARCPHTRC 191

RESULT 6
OX40_HUMAN
ID OX40_HUMAN STANDARD; PRT; 277 AA.
AC P43489; Q13663;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE OX40L RECEPTOR PRECURSOR (ACT35 ANTIGEN) (TAX-TRANSCRIPTIONALLY
DE ACTIVATED GLYCOPROTEIN 1 RECEPTOR) (CD134 ANTIGEN).
GN TNFRSF4 OR TXGP1L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94170844; PubMed=7510240;
RA Latza U., Duerkop H., Schnittger S., Ringeling J., Eitelbach F.,
RA Hummel M., Fonatsch C., Stein H.;
RT "The human OX40 homolog: cDNA structure, expression and chromosomal
RT assignment of the ACT35 antigen."
RL Eur. J. Immunol. 24:677-683(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95219871; PubMed=7704935;
RA Baum P.R., Gayle R.B. III, Ramsdell F., Srinivasan S., Sorensen R.A.,
RA Watson M.L., Seldin M.F., Clifford K.N., Grabstein K., Alderson M.R.;
RT "Identification of OX40 ligand and preliminary characterization of
RT its activities on OX40 receptor."
RL Circ. Shock 44:30-34(1994).
CC -!- FUNCTION: RECEPTOR FOR THE OX40L/GP34 CYTOKINE.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD134 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd134.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X75962; CAA53576.1; -
DR EMBL; S76792; AAB33944.1; ALT_INIT.
DR HSP; P25942; 1CDF.
DR MIM; 600315; -
DR INTERPRO; IPR001368; -
DR PFAM; PF00020; TNFR_C6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; 3.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
```

```
KW Receptor; T-cell; Antigen; Glycoprotein; Transmembrane; Repeat;
KW Signal.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 277 OX40L RECEPTOR.
FT DOMAIN 29 214 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 215 235 POTENTIAL.
FT DOMAIN 236 277 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 30 169 4 X TNFR-CYS.
FT REPEAT 30 65 TNFR-CYS 1.
FT REPEAT 66 107 TNFR-CYS 2.
FT REPEAT 108 126 TNFR-CYS 3 (INCOMPLETE).
FT REPEAT 127 167 TNFR-CYS 4.
FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 277 AA; 29340 MW; 49F15525941550BF CRC64;

Query Match 36.6%; Score 93; DB 1; Length 277;
Best Local Similarity 45.5%; Pred. No. 0.0075;
Matches 20; Conservative 7; Mismatches 13; Indels 4; Gaps 2;

QY 1 CPPGAGVIAPGTPSQNTQCPCPPGTFSSSSSSSQCCQPHRNC 44
||||| : ||||| : ||||| : ||||| : ||||| :
Db 46 CRPGMGVSRCSRSQNTVCRPGPG-FYNDVVSCKPKP---CT 85

RESULT 7
TNRC_MOUSE
ID TNRC_MOUSE STANDARD; PRT; 415 AA.
AC P50284;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE LYMPHOTOXIN-BETA RECEPTOR PRECURSOR.
GN LTBR OR TNFR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CVB; TISSUE=LUNG;
RX MEDLINE=96072804; PubMed=7594541;
RA Force W.R., Walter B.N., Hession C., Tizard R., Kozak C.A.,
RA Browning J.L., Ware C.F.;
RT "Mouse lymphotoxin-beta receptor. Molecular genetics, ligand binding,
RT and expression."
RL J. Immunol. 155:5280-5288(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96163885; PubMed=8586432;
RA Nakamura T., Tashiro K., Nazarea M., Nakano T., Sasayama S.,
RA Honjo T.;
RT "The murine lymphotoxin-beta receptor cDNA: isolation by the signal
RT sequence trap and chromosomal mapping."
RL Genomics 30:312-319(1995).
CC -!- FUNCTION: RECEPTOR FOR THE LYMPHOTOXIN-BETA. POSSIBLE FUNCTION IN
CC IMMUNE DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U29173; AAA6864.1; -
DR EMBL; U38423; AAB00846.1; -
DR EMBL; U30798; AAA81334.1; -
DR HSP; P25942; 1CDF.
DR MGD; MGI:104875; LTBR.
```


RESULT	12
NGFR_RAT	
ID	NGFR_
AC	P0717

```
FT DISULFID 33 44 BY SIMILARITY.
FT DISULFID 45 58 BY SIMILARITY.
FT DISULFID 48 65 BY SIMILARITY.
FT DISULFID 68 84 BY SIMILARITY.
FT DISULFID 87 100 BY SIMILARITY.
FT DISULFID 90 108 BY SIMILARITY.
FT DISULFID 110 123 BY SIMILARITY.
FT DISULFID 126 139 BY SIMILARITY.
FT DISULFID 129 147 BY SIMILARITY.
FT DISULFID 150 165 BY SIMILARITY.
FT DISULFID 168 181 BY SIMILARITY.
FT DISULFID 171 189 BY SIMILARITY.
FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 71 71 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 425 AA; 45432 MW; B2E152D94D3827F8 CRC64;

Query Match 34.1%; Score 86.5; DB 1; Length 425;
Best Local Similarity 36.4%; Pred. No. 0.048;
Matches 16; Conservative 8; Mismatches 17; Indels 3; Gaps 1;

Qy 1 CPPGAGVTAQPTSONTCQCPPTGTFSSSSSSQCOQPHRNCT 44
Db 129 CEVSGSLVFCQDKQNTVCECEPCTGYSDANHYDPCLP---CT 169

RESULT 13
OX40_MOUSE
ID OX40_MOUSE STANDARD; PRT; 272 AA.
AC P47741;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE OX40L RECEPTOR PRECURSOR (OX40 ANTIGEN).
GN TNFRSF4 OR TKG1 OR OX40.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE-94044750; PubMed=8228223;
RA Calderhead D.M., Buhlmann J.E., van den Eertwegh A.J.,
RA Claassen E., Noelle R.J., Fell H.;
RT "Cloning of mouse OX40; a T cell activation marker that may mediate
RT T-B cell interactions."
RL J. Immunol. 151:5261-5271(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-95255413; PubMed=7737295;
RA Birkeland M.L., Copeland N.G., Gilbert D.J., Jenkins N.A.,
RA Barclay A.N.;
RT "Gene structure and chromosomal localization of the mouse homologue
RT of rat OX40 protein."
RL Eur. J. Immunol. 25:926-930(1995).
CC -!- FUNCTION: RECEPTOR FOR THE OX40L/GP34 CYTOKINE.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z21674; CAA79772.1; -
CC EMBL: X85214; CAA59476.1; -
CC DR HSSP: P25942; ICDF.
CC DR MGD: MGI:104512; TXGP1.
CC DR INTERPRO: IPR001368; -
CC PFAM: PF00020; TNFR_c6; 3.
```

```
DR PROSITE; PS00652; TNFR_NGFR_1; 3.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
KW Receptor; T-cell; Antigen; Glycoprotein; Transmembrane; Repeat;
KW Signal.
FT CHAIN 1 19 POTENTIAL.
FT CHAIN 20 272 OX40L RECEPTOR.
FT DOMAIN 20 211 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 212 236 POTENTIAL.
FT DOMAIN 237 272 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 26 165 4 X TNFR-CYS.
FT REPEAT 26 61 TNFR-CYS 1.
FT REPEAT 26 61 TNFR-CYS 2.
FT REPEAT 104 124 TNFR-CYS 3 (INCOMPLETE).
FT REPEAT 125 165 TNFR-CYS 4.
FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 15 15 A -> G (IN REF. 2).
SQ SEQUENCE 272 AA; 30153 MW; 06E7BB4156F0D08E CRC64;

Query Match 33.9%; Score 86; DB 1; Length 272;
Best Local Similarity 37.0%; Pred. No. 0.037;
Matches 20; Conservative 7; Mismatches 15; Indels 12; Gaps 4;

Qy 1 CPPGAGVIA---PGT-PSQNT-----QCQPCPPGTFSSSSSSQCOQPHRNCT 44
Db 95 CTPTQDTVCRRCRPGTQPRQDSGYKLGVDVCPGPGHF--SPGNNAACKPWNTCT 146

RESULT 14
VT2_SFVKA
ID VT2_SFVKA STANDARD; PRT; 325 AA.
AC P25943;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE TUMOR NECROSIS FACTOR SOLUBLE RECEPTOR PRECURSOR (PROTEIN T2).
GN T2.
OS Shope fibroma virus (strain kasza) (SFV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Leporipoxvirus.
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE-87321103; PubMed=2820128;
RA Upton C., Delange A.M., McFadden G.;
RT "Tumorigenic poxviruses: genomic organization and DNA sequence of the
RT telomeric region of the Shope fibroma virus genome."
RL Virology 160:20-30(1987).
RN [2]
RP FUNCTION.
RX MEDLINE-91207415; PubMed=1850261;
RA Smith C.A., Davis T., Wignall J.M., Din W.S., Farrah T., Upton C.,
RA McFadden G., Goodwin R.G.;
RT "T2 open reading frame from the Shope fibroma virus encodes a soluble
RT form of the TNF receptor."
RL Biochem. Biophys. Res. Commun. 176:335-342(1991).
CC -!- FUNCTION: BINDS TO TNF-ALPHA AND BETA. PROBABLY PREVENTS TNF TO
CC REACH CELLULAR TARGET AND THEREBY DEAMPENING THE POTENTIAL
CC ANTIVIRAL EFFECTS OF THE CYTOKINE.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M17433; -; NOT_ANNOTATED_CDS.
CC EMBL: A23727; CAA01687.1; -
CC DR PIR: B43692; B43692.
CC DR HSSP: P19438; 1TNR.
CC DR INTERPRO: IPR001368; -
```

ONLY EXPRESSED IN ALVEOLAR MACROPHAGES IS PRODUCED BY THE USE OF AN ALTERNATIVE INITIATION CODON IN THE SAME READING FRAME. -1- PTM: PHOSPHORYLATED ON SERINE AND TYROSINE RESIDUES. -1- DISEASE: MOST SPECIFIC HODGKIN'S DISEASE ASSOCIATED ANTIGEN. -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION. -1- DATABASE: NAME-PROW; NOTE-CD guide CD30 entry; WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd30.htm".

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; M83554; AAA51947.1; -
DR EMBL; S75768; AAD14188.1; -
DR EMBL; D86042; BAA12973.1; -
DR PIR; A42086; A42086.
DR HSSP; P19438; INCF.
DR MIM; 153243; -
DR INTERPRO; IPR001368; -
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00650; TNFR_NGFR_2; 2.
DR T-cell; Glycoprotein; Transmembrane; Signal; Receptor; Repeat;
KW Phosphorylation; Alternative initiation.
FT SIGNAL 1 18
FT CHAIN 19 595 CD30L RECEPTOR.
FT CHAIN 464 595 CD30L RECEPTOR, VARIANT C30V.
FT INIT_MET 464 464 FOR VARIANT C30V.
FT DOMAIN 19 379 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 380 407 POTENTIAL.
FT DOMAIN 408 595 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 28 325 6 X TNFR-CYS.
FT REPEAT 28 66 TNFR-CYS 1.
FT REPEAT 68 106 TNFR-CYS 2.
FT REPEAT 107 150 TNFR-CYS 3.
FT REPEAT 205 241 TNFR-CYS 4.
FT REPEAT 243 281 TNFR-CYS 5.
FT REPEAT 282 325 TNFR-CYS 6.
FT DOMAIN 347 377 PRO/SER/THR-RICH.
FT CARBOHYD 101 101 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 595 AA; 63747 MW; 7A407CC78A6E0BC8 CRC64;

Query Match 32.9%; Score 83.5; DB 1; Length 595;
Best Local Similarity 44.7%; Pred. No. 0.13;
Matches 17; Conservative 5; Mismatches 15; Indels 1; Gaps 1;

Qy 1 CPPGAGVIAPCTPSQNTQCPCPGTFSASSSSSEQC 38
Db 131 CPAGMIVKFPCTAGKNTVCPEPSPGV-SPACASPEACK 167

Search completed: January 30, 2001, 17:05:57
Job time: 1218 sec

PFAM; PF00020; TNFR_C6; 2.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00650; TNFR_NGFR_2; 1.
DR Receptor; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 325 TUMOR NECROSIS FACTOR SOLUBLE RECEPTOR.
FT DOMAIN 27 186 4 X TNFR-CYS.
FT REPEAT 27 62 TNFR-CYS 1.
FT REPEAT 63 104 TNFR-CYS 2.
FT REPEAT 105 147 TNFR-CYS 3.
FT REPEAT 148 186 TNFR-CYS 4.
FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 325 AA; 35132 MW; 810530339198A71E CRC64;

Query Match 32.9%; Score 83.5; DB 1; Length 325;
Best Local Similarity 45.9%; Pred. No. 0.076;
Matches 17; Conservative 7; Mismatches 12; Indels 1; Gaps 1;

Qy 1 CPPGAGVIAPCTPSQNTQCPCPGTFSASSSSSEQC 37
Db 129 CPAGVGV-SGTHRAGDTLCEKCPHTYSDLSLPTERC 164

RESULT 15
CD30_HUMAN
ID CD30_HUMAN STANDARD; PRT; 595 AA.
AC P28908;
DT 01-DEC-1992 (Rel. 24, Created)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CD30L RECEPTOR PRECURSOR (LYMPHOCYTE ACTIVATION ANTIGEN CD30) (KI-1 ANTIGEN).
GN TNFRSF8 OR CD30.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPHOID;
RX MEDLINE=92154659; PubMed=1310894;
RA Duerkop H., Latza U., Hummel M., Eitelbach F., Seed B., Stein H.;
RT "Molecular cloning and expression of a new member of the nerve growth factor receptor family that is characteristic for Hodgkin's disease.";
RT Cell 68:421-427(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95089787; PubMed=7527901;
RA Jung W., Krueger S., Renner C., Gause A., Sahin U., Trumper L., Pfrendrich M.;
RT "Opposite effects of the CD30 ligand are not due to CD30 mutations: results from cDNA cloning and sequence comparison of the CD30 antigen from different sources.";
RT Mol. Immunol. 31:1329-1334(1994).
RN [3]
RP SEQUENCE FROM N.A. (VARIANT C30V).
RX MEDLINE=96437016; PubMed=8839832;
RA Horie R., Ito K., Tatewaki M., Nagai M., Aizawa S., Higashihara M., Ishida T., Inoue J., Takizawa H., Watanabe T.;
RT "A variant CD30 protein lacking extracellular and transmembrane domain is induced in HL-60 by tetradecanoylphorbol acetate and is expressed in alveolar macrophages.";
RT Blood 88:2422-2432(1996).
CC -1- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD30L. MAY PLAY A ROLE IN THE REGULATION OF CELLULAR GROWTH AND TRANSFORMATION OF ACTIVATED LYMPHOBLASTS. REGULATES GENE EXPRESSION THROUGH ACTIVATION OF NFkB.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- ALTERNATIVE PRODUCTS: A SHORTER CYTOPLASMIC FORM C30V WHICH IS

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model
Run on: January 30, 2001, 16:54:41 ; Search time 254.27 Seconds
(without alignments)
20.282 Million cell updates/sec

Title: US-09-518-931-2_COPY_132_175
Perfect score: 254
Sequence: 1 CPPGAGVIAPGPSQNTQCO.....GTFSSSSSSSQCPHRNCT 44

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_15:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phage:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	254	100.0	300	4	O95407
2	142	55.9	401	11	O08712
3	142	55.9	401	11	O08727
4	140	55.1	372	4	O0UHF4
5	140	55.1	401	4	O00300
6	136	53.5	439	4	O16042
7	126	49.6	459	11	O62327
8	126	49.6	482	11	O08734
9	123.5	48.6	430	6	O9N092
10	116	45.7	107	11	O9JKE0
11	107.5	42.3	283	6	O9XS28
12	107	42.1	302	13	O9PU50
13	105	41.3	616	4	O9Y606
14	103.5	40.7	283	4	O92956
15	103.5	40.7	283	4	O9UM65
16	103	40.6	175	11	O9WUL4
17	97	38.2	625	11	O35305
18	94.5	37.2	142	13	O91932
19	93	36.6	655	4	O75509

20	92	36.2	186	12	O72735
21	91	35.8	30	4	O9UIH1
22	90	35.4	234	4	O95851
23	90	35.4	241	4	O9Y505
24	89.5	35.2	143	13	O919J3
25	89.5	35.2	401	13	O9PRG7
26	88	34.6	186	12	O9WJBA
27	87	34.3	186	12	O9YB87
28	86.5	34.1	387	13	O9PVD4
29	85.5	33.7	417	11	O9ZOW1
30	85	33.5	255	4	O9NKJ9
31	84.5	33.3	267	6	O02764
32	79.5	31.3	348	12	O57277
33	79.5	31.3	348	12	O57103
34	79.5	31.3	348	12	O57108
35	79.5	31.3	349	12	O57291
36	79.5	31.3	349	12	O57099
37	79.5	31.3	349	12	O57100
38	79.5	31.3	349	12	O57101
39	79.5	31.3	349	12	O57102
40	79.5	31.3	360	12	O57118
41	79.5	31.3	1019	5	O9NA40
42	79	31.1	326	12	O57120
43	79	31.1	326	12	O57122
44	79	31.1	349	12	O57305
45	79	31.1	351	12	O57117

ALIGNMENTS

RESULT 1

O95407

ID O95407 PRELIMINARY; PRT; 300 AA.

AC O95407;

DT 01-MAY-1999 (TrEMBLrel. 10, Created)

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)

DE DECOY RECEPTOR 3 (M68) (M68C) (M68E).

GN DCR3 OR TR6.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCBI_TaxID=9606;

EN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99087326; PubMed=9872321;

RA Pitti R.M., Marsters S.A., Lawrence D.A., Roy M., Kischkel F.C.,

RA Dowd P., Huang A., Donahue C.J., Sherwood S.W., Baldwin D.T.,

RA Godowski P.J., Wood W.I., Gurney A.L., Hillan K.J., Cohen R.L.,

RA Goddard A.D., Botstein D., Ashkenazi A.;

RT "Genomic amplification of a decoy receptor for Fas ligand in lung and

RT colon cancer.";

RL Nature 396:699-703(1998).

RN [2]

RP SEQUENCE FROM N.A.

EC TISSUE=BLOOD;

RA	YU K.Y., Kwon B., Ni J., Zhai Y., Ebner R., Kwon B.S.;
RT	"A newly identified member of tumor necrosis factor receptor
RT	superfamily (TR6) suppresses LIGHT-mediated apoptosis.";
RL	J. Biol. Chem. 274:13733-13736(1999).
RN [3]	
RP	SEQUENCE FROM N.A.
RC	TISSUE=PANCREAS;
RX	MEDLINE=20123600; PubMed=10655513;
RA	Bai C., Connolly B., Metzker M.L., Hilliard C.A., Liu X., Sandig V.,
RA	Soderman A., Galloway S.M., Liu Q., Austin C.P., Caskey C.T.;
RT	"Overexpression of M68/DCR3 in human gastrointestinal tract tumors
RT	independent of gene amplification and its location in a four-gene
RT	cluster.";
RL	Proc. Natl. Acad. Sci. U.S.A. 97:1230-1235(2000).
DR	EMBL; AF104419; AAD03056.1; -.

15 TO DAY 17.

- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.

CC EMBL; AF134240; AAD29688.1; -

CC EMBL; AF217796; AAF35244.1; -

CC EMBL; AF217793; AAF33685.1; -

CC EMBL; AF217794; AAF33686.1; -

CC HSSP; P25942; ICDF.

CC INTERPRO; IPR000561; -

CC INTERPRO; IPR001368; -

CC PFAM; PF00020; TNFR_C6; 4.

CC PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.

CC PROSITE; PS01186; EGF_2; UNKNOWN_1.

CC PROSITE; PS00500; TNFR_NGFR_2; 2.

CC PRODOM; PD000771; -; 1.

CC Receptor.

CC KW Receptor.

CC SQ SEQUENCE 300 AA; 32679 MW; F90AEE33718449AF CRC64;

Query Match 100.0%; Score 254; DB 4; Length 300;

Best Local Similarity 100.0%; Pred. No. 6.5e-23;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CPPGAGVATPQNTQCPCPPTGTFSSSSSSSQCPHNRCT 44

DB 132 CPPGAGVATPQNTQCPCPPTGTFSSSSSSSQCPHNRCT 175

RESULT 2

008712 PRELIMINARY; PRT; 401 AA.

ID 008712; 070202;

AC 01-JUL-1997 (TrEMBLrel. 04, Created)

DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)

DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)

DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR) (OCIF)

GN TNFRSF11B OR OPG.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OC NCBI_Taxid=10090;

OX [1]

RN SEQUENCE FROM N.A.

RP STRAIN=BALB/C; TISSUE=KIDNEY;

RX MEDLINE=97262071; PubMed=9108485;

RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S., Luethy R., Nguyen H.Q., Wooden S., Bennett L., Boone T., Shimamoto G., Daro E., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J., Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W., Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R., Suggs S., Boyle W.J.

RA "Osteoprotegerin: a novel secreted protein involved in the regulation of bone density."

RT Cell 89:309-319(1997).

RL [2]

RN SEQUENCE FROM N.A.

RP STRAIN=129/OLA, AND NIH SWISS;

RX MEDLINE=98382527; PubMed=9714833;

RA Mizuno A., Murakami A., Nakagawa N., Yasuda H., Tsuda E., Morinaga T., Higashio K.

RA "Structure of the mouse osteoclastogenesis inhibitory factor (OCIF) gene and its expression in embryogenesis."

RT Gene 215:339-343(1998).

RL -1- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN STROMAL CELLS AND OSTEOCLAST PROGENITORS.

CC -1- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM).

CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.

CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN LIVER, LUNG, BRAIN, HEART, KIDNEY, STOMACH, INTESTINE, SKIN, CALVARIA AND PLACENTA. NOT DETECTED IN SPLEEN.

CC -1- DEVELOPMENTAL STAGE: IN THE EMBRYO, HIGH LEVELS ARE DETECTED AT EUTEROTIC STAGE. EXPRESSION DECREASES AND THEN INCREASES FROM DAY 15 TO DAY 17.

CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.

CC EMBL; AF134240; AAD29688.1; -

CC EMBL; AF217796; AAF35244.1; -

CC EMBL; AF217793; AAF33685.1; -

CC EMBL; AF217794; AAF33686.1; -

CC HSSP; P25942; ICDF.

CC INTERPRO; IPR000561; -

CC INTERPRO; IPR001368; -

CC PFAM; PF00020; TNFR_C6; 4.

CC PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.

CC PROSITE; PS01186; EGF_2; UNKNOWN_1.

CC PROSITE; PS00500; TNFR_NGFR_2; 2.

CC PRODOM; PD000771; -; 1.

CC Receptor.

CC KW Receptor.

CC SQ SEQUENCE 300 AA; 32679 MW; F90AEE33718449AF CRC64;

Query Match 100.0%; Score 254; DB 4; Length 300;

Best Local Similarity 100.0%; Pred. No. 6.5e-23;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CPPGAGVATPQNTQCPCPPTGTFSSSSSSSQCPHNRCT 44

DB 132 CPPGAGVATPQNTQCPCPPTGTFSSSSSSSQCPHNRCT 175

RESULT 2

008712 PRELIMINARY; PRT; 401 AA.

ID 008712; 070202;

AC 01-JUL-1997 (TrEMBLrel. 04, Created)

DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)

DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)

DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR) (OCIF)

GN TNFRSF11B OR OPG.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OC NCBI_Taxid=10090;

OX [1]

RN SEQUENCE FROM N.A.

RP STRAIN=BALB/C; TISSUE=KIDNEY;

RX MEDLINE=97262071; PubMed=9108485;

RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S., Luethy R., Nguyen H.Q., Wooden S., Bennett L., Boone T., Shimamoto G., Daro E., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J., Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W., Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R., Suggs S., Boyle W.J.

RA "Osteoprotegerin: a novel secreted protein involved in the regulation of bone density."

RT Cell 89:309-319(1997).

RL [2]

RN SEQUENCE FROM N.A.

RP STRAIN=129/OLA, AND NIH SWISS;

RX MEDLINE=98382527; PubMed=9714833;

RA Mizuno A., Murakami A., Nakagawa N., Yasuda H., Tsuda E., Morinaga T., Higashio K.

RA "Structure of the mouse osteoclastogenesis inhibitory factor (OCIF) gene and its expression in embryogenesis."

RT Gene 215:339-343(1998).

RL -1- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN STROMAL CELLS AND OSTEOCLAST PROGENITORS.

CC -1- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM).

CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.

CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN LIVER, LUNG, BRAIN, HEART, KIDNEY, STOMACH, INTESTINE, SKIN, CALVARIA AND PLACENTA. NOT DETECTED IN SPLEEN.

CC -1- DEVELOPMENTAL STAGE: IN THE EMBRYO, HIGH LEVELS ARE DETECTED AT EUTEROTIC STAGE. EXPRESSION DECREASES AND THEN INCREASES FROM DAY 15 TO DAY 17.

CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.

CC EMBL; AF134240; AAD29688.1; -

CC EMBL; AF217796; AAF35244.1; -

CC EMBL; AF217793; AAF33685.1; -

CC EMBL; AF217794; AAF33686.1; -

CC HSSP; P25942; ICDF.

CC INTERPRO; IPR000561; -

CC INTERPRO; IPR001368; -

CC PFAM; PF00020; TNFR_C6; 4.

CC PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.

CC PROSITE; PS01186; EGF_2; UNKNOWN_1.

CC PROSITE; PS00500; TNFR_NGFR_2; 2.

CC PRODOM; PD000771; -; 1.

CC Receptor.

CC KW Receptor.

CC SQ SEQUENCE 300 AA; 32679 MW; F90AEE33718449AF CRC64;

Query Match 100.0%; Score 254; DB 4; Length 300;

Best Local Similarity 100.0%; Pred. No. 1.8e-09;

Matches 23; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 1 CPPGAGVATPQNTQCPCPPTGTFSSSSSSSQCPHNRCT 44

DB 124 CPPGAGVATPQNTQCPCPPTGTFSSSSSSSQCPHNRCT 167

RESULT 3

008727 PRELIMINARY; PRT; 401 AA.

ID 008727

AC 008727;

DT 01-JUL-1997 (TrEMBLrel. 04, Created)

DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)

DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)

DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR) (OCIF)

GN TNFRSF11B OR OPG.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
TISSUE=PLACENTA;
He Z.-Y., Yang G.-Z., Zhang W.-J., Wu X.-F.;
"Cloning and Expression of Osteoprotegerin from Homo sapiens.";
Sheng Wu Hua Hsueh Yu Sheng Wu Li Hsueh Pao 31:680-684(1999).
EMBL; AF134187; AAF20168.1; -;
HSP; P25942; 1CDF.
DR INTERPRO; IPR001368; -;
DR PFAM; PF00020; TNFR_C6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
FT NON_TER 1
SQ SEQUENCE 372 AA; 42758 MW; F02527A5CD01CCD3 CRC64;

Query Match 55.1%; Score 140; DB 4; Length 372;
Best Local Similarity 52.3%; Pred.No. 2.9e-09;
Matches 23; Conservative 6; Mismatches 15; Indels 0; Gaps

Qy 1 CPFGAGVIAPGTSTONTQCPCPPGTFSSASSSEOCOPHRNCT 44
|||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||:
Db 103 CPFGGVVQAATPERNTVCKRCPDGFFSNETSSKAPCRHTNCS 146

RESULT 5
O00300 PRELIMINARY; PRT; 401 AA.

ID O00300 PRELIMINARY;
AC O00300; O60236;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE OSTEOPTERGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR) (OCIF) (TUMOR NECROSIS FACTOR RECEPTOR SUPERFAMILY MEMBER 11B).
GN TNFRSF11B OR OPB OR OCIF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=KIDNEY;
RC MEDLINE=97262071; PubMed=9108485;
RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S., Luethy R., Nguyen H.O., Wooden S., Bennett L., Boone T., Shimamoto G.A.
RA Derose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.J.
RA Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W., Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R., Suggs S., Boyle W.J.;
RT "Osteoprotegerin: a novel secreted protein involved in the regulation of bone density";
RL Cell 89:309-319(1997).
[2]
RN SEQUENCE FROM N.A.
RP TISSUE=LUNG FIBROBLAST;
RC MEDLINE=98151033; PubMed=9492069;
RA Yasuda H., Shima N., Nakagawa N., Mochizuki S.-I., Yano K., Fujise N., Sato Y., Goto M., Yamaguchi K., Kuriyama M., Kanno T., Murakami A., Tsuda E., Morinaga T., Higashio K.;
RT "Identity of osteoclastogenesis inhibitory factor (OCIF) and osteoprotegerin (OPG): a mechanism by which OPG/OCIF inhibits osteoclastogenesis in vitro.";
RL Endocrinology 139:1329-1337(1998).
[3]
RN SEQUENCE FROM N.A.
RP TISSUE=PLACENTA;
RC MEDLINE=98351569; PubMed=9688283;
RA Morinaga T., Nakagawa N., Yasuda H., Tsuda E., Higashio K.;
RT "Cloning and characterization of the gene encoding human osteoprotegerin/osteoclastogenesis-inhibitory factor.";
RL Eur. J. Biochem. 254:685-691(1998).
CC "- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES

RA Dembic Z., Loetscher H., Gudger U., Paul T.C., Laitin H.W., Schütz R. 1977

DT 01-NOV-1998 (TREMBlé, 08, Created)

RA Dembic Z., Loetscher H., Gudger U., Paul T.C., Laitin H.W., Schütz R. 1977


```

CC      3 PGAGVIAGTSPQNTQCPCPPGTFSASSSSSEQCQPHRNC 44
CC      EMBL; AF018253; AAB86809.1; -.
CC      HSSP; P25942; 1CDF.
DB      146 PGQVQKGTESQDTLCQNCPPGTFS-SNGTLECCQHGNKS 186

RESULT 12
Q9PUS0 PRELIMINARY; PRT; 302 AA.
AC Q9PUS0
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE DECOY RECEPTOR.
OS Salvelinus fontinalis (Brook trout).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
OX NCBI_TaxID=8038;
RN [1]
RP SEQUENCE FROM N.A.
RA Boe J., Goetz F.W.;
RT "A tumor necrosis factor receptor homolog is up-regulated in the brook
RT trout (Salvelinus fontinalis) ovary at the completion of ovulation.";
RL Biol. Reprod. 0:0-0(1999).
DR EMBL; AF156738; AAD56428.1; -.
DR HSSP; P19438; 1EXT.
DR INTERPRO; IPR000561; -.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS50050; TNFR_NGFR_2; 1.
KW Receptor.
SQ SEQUENCE 302 AA; 34037 MW; E44C73477F05C3DF CRC64;

Query Match 42.1%; Score 107; DB 13; Length 302;
Best Local Similarity 44.2%; Pred. No. 2e-05;
Matches 19; Conservative 4; Mismatches 20; Indels 0; Gaps 0;

QY 1 CPPAGVIAGTSPQNTQCPCPPGTFSASSSSSEQCQPHRNC 43
DB 120 CPPGYGANTTGTPHQTECVQCQAGFYSEVSSAKATCLAOSNC 162

RESULT 13
Q9Y6Q6 PRELIMINARY; PRT; 616 AA.
AC Q9Y6Q6
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE RECEPTOR ACTIVATOR OF NF-KAPPA-B PRECURSOR (TNF-RELATED ACTIVATION-
DE INDUCED CYTOKINE RECEPTOR) (RANK).
GN TNFRSF11A OR RANK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Anderson D.M., Maraskovsky E., Billingsley W.L., Dougall W.C.,
RA Tometsko M.E., Roux E.R., Teepe M.C., DuBose R.F., Cosman D.,
RA Galibert L.;
RT "A homologue of the TNF receptor and its ligand enhance T-cell growth
RT and dendritic-cell function.";
RL Nature 390:175-179(1997).
CC -!- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS RANKL.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).
CC -!- TISSUE SPECIFICITY: UBIQUITOUS EXPRESSION WITH HIGH LEVELS IN
CC SKELETAL MUSCLE, THYMUS, LIVER, COLON, SMALL INTESTINE AND ADRENAL

```

```

CC      1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC      EMBL; AF018253; AAB86809.1; -.
CC      HSSP; P25942; 1CDF.
CC      MIM; 603499; -.
DB      INTERPRO; IPR001368; -.
DB      PFAM; PF00020; TNFR_C6; 4.
DB      PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DB      PROSITE; PS50050; TNFR_NGFR_2; 1.
KW Receptor; Glycoprotein; Transmembrane; Signal; Repeat.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 616 RECEPTOR ACTIVATOR OF NF-KAPPA-B.
FT DOMAIN 24 212 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 213 233 POTENTIAL.
FT DOMAIN 234 616 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 33 195 4 X TNFR-CYS.
FT REPEAT 33 69 TNFR-CYS 1.
FT REPEAT 70 112 TNFR-CYS 2.
FT REPEAT 113 152 TNFR-CYS 3.
FT REPEAT 153 195 TNFR-CYS 4.
FT DISULFID 34 46 BY SIMILARITY.
FT DISULFID 47 60 BY SIMILARITY.
FT DISULFID 50 68 BY SIMILARITY.
FT DISULFID 71 86 BY SIMILARITY.
FT DISULFID 92 112 BY SIMILARITY.
FT DISULFID 114 124 BY SIMILARITY.
FT DISULFID 126 133 BY SIMILARITY.
FT DISULFID 127 151 BY SIMILARITY.
FT DISULFID 154 169 BY SIMILARITY.
FT DISULFID 175 194 BY SIMILARITY.
FT CARBOHYD 105 105 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 174 174 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 616 AA; 66033 MW; E3DE9A7A08196F81 CRC64;

Query Match 41.3%; Score 105; DB 4; Length 616;
Best Local Similarity 43.2%; Pred. No. 6.5e-05;
Matches 19; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

QY 1 CPPAGVIAGTSPQNTQCPCPPGTFSASSSSSEQCQPHRNC 44
DB 133 CAPGLGAQHPLQLNKDTCVCKPLAGYFSDAFSTDKRPTWNC 176

RESULT 14
Q92956 PRELIMINARY; PRT; 283 AA.
AC Q92956
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE HERPESVIRUS ENTRY MEDIATOR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Montgomery R.I., Warner M.S., Lum B.J., Spear P.G.;
RT "Herpes simplex virus-1 entry into cells mediated by a novel member of
RT the TNF/NGF receptor family.";
RL Cell 87:427-436(1996).
DR EMBL; U70321; AAB58354.1; -.
DR HSSP; P25942; 1CDF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -.
SQ SEQUENCE 283 AA; 30420 MW; DFCE146E4E024F4B CRC64;

```


THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:47:41 ; Search time 132.77 seconds
(without alignments)
5.951 Million cell updates/sec

Title: US-09-518-931-2_COPY_132_175

Perfect score: 254

Sequence: 1 CPPGAGVIAPGTPSQNTQCO.....GTFSSSSSSSEQCPHRNCT 44

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA.*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/5_COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	254	100.0	300	2	US-08-794-796-2
2	142	55.9	401	3	US-08-974-022-2
3	142	55.9	401	3	US-08-974-022-4
4	140	55.1	401	3	US-08-974-022-6
5	136	53.5	163	2	US-08-219-237B-5
6	136	53.5	164	2	US-08-232-087A-9
7	136	53.5	227	3	US-08-974-022-48
8	136	53.5	461	1	US-08-385-229-2
9	136	53.5	461	2	US-08-650-000-2
10	136	53.5	461	5	5395760-2
11	136	53.5	486	1	US-08-243-010-1
12	136	53.5	518	1	US-08-385-229-4
13	126	49.6	474	2	US-08-650-000-4
14	126	49.6	474	5	5395760-4
15	107	42.1	197	3	US-08-974-022-49
16	105	41.3	451	3	US-08-996-139-4
17	105	41.3	591	3	US-08-996-139-2
18	105	41.3	616	3	US-08-996-139-6
19	103.5	40.7	283	4	PCT-US96-12374-2
20	97	38.2	162	2	US-08-219-237B-7
21	97	38.2	625	3	US-08-996-139-15
22	94.5	37.2	197	2	US-08-505-606-1
23	93	36.6	277	2	US-08-147-784-2
24	93	36.6	555	3	US-08-959-382-2
25	90	35.4	228	3	US-08-911-423-6
26	90	35.4	241	3	US-08-911-423-4
27	90	35.4	311	3	US-08-911-423-8
28	89	35.0	139	2	US-08-219-237B-8

29	89	35.0	205	3	US-08-974-022-51	Sequence 51, Appl
30	87.5	34.4	159	2	US-08-219-237B-6	Sequence 6, Appl
31	87.5	34.4	224	3	US-08-974-022-50	Sequence 50, Appl
32	87	34.3	186	1	US-08-089-458B-6	Sequence 6, Appl
33	86.5	34.1	159	2	US-08-232-087A-11	Sequence 11, Appl
34	86	33.9	206	1	US-08-097-827-7	Sequence 7, Appl
35	86	33.9	206	1	US-08-494-574-7	Sequence 7, Appl
36	86	33.9	438	1	US-08-494-574-7	Sequence 11, Appl
37	86	33.9	438	1	US-08-494-574-11	Sequence 11, Appl
38	83.5	32.9	207	3	US-08-974-022-47	Sequence 47, Appl
39	83.5	32.9	325	1	US-08-292-549-2	Sequence 2, Appl
40	83.5	32.9	325	4	PCT-US91-02207-2	Sequence 2, Appl
41	83.5	32.9	595	1	US-08-225-989-2	Sequence 2, Appl
42	83.5	32.9	595	1	US-08-570-923-2	Sequence 2, Appl
43	83.5	32.9	595	1	US-08-580-014-2	Sequence 2, Appl
44	83.5	32.9	595	2	US-08-232-087A-2	Sequence 2, Appl
45	83.5	32.9	595	3	US-09-079-785-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-794-796-2
; Sequence 2, Application US/08794796
; Patent No. 5885800
; GENERAL INFORMATION:
; APPLICANT: Emery, John
; APPLICANT: Tan, KB
; APPLICANT: Truneh, Alem
; APPLICANT: Young, Peter
; TITLE OF INVENTION: Tumor Necrosis Related Receptor,
; TITLE OF INVENTION: TR4
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/794,796
; FILING DATE: 04-FEB-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: GH50000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-4026
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 300 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-794-796-2

Query Match 100.0%; Score 254; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 3.9e-19;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CPPGAGVIAGTPTSQNTQCPCPPGTFSASSSSSEQCQPHRNCT 44
|||||
Db 132 CPPGAGVIAGTPTSQNTQCPCPPGTFSASSSSSEQCQPHRNCT 175
|||||

RESULT 2
US-08-974-022-2
; Sequence 2, Application US/08974022
; Patent No. 6015938
; GENERAL INFORMATION:
; APPLICANT: Boyle, William J.
; APPLICANT: Lacey, David L.
; APPLICANT: Calzone, Frank J.
; APPLICANT: Chang, Ming-Shi
; TITLE OF INVENTION: OSTEOPROTEGERIN
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,022
; FILING DATE: 12-DEC-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/577,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-378
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 401 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-974-022-4

Query Match 55.9%; Score 142; DB 3; Length 401;
Best Local Similarity 52.3%; Pred. No. 1.4e-07;
Matches 23; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 1 CPPGAGVIAGTPTSQNTQCPCPPGTFSASSSSSEQCQPHRNCT 44
|||||
Db 124 CPPGSGVQAGTPTERTVCKPCDGFSGTSSKAPCKKHTNCS 167
|||||

RESULT 4
US-08-974-022-6
; Sequence 6, Application US/08974022
; Patent No. 6015938
; GENERAL INFORMATION:
; APPLICANT: Boyle, William J.
; APPLICANT: Lacey, David L.
; APPLICANT: Calzone, Frank J.
; APPLICANT: Chang, Ming-Shi
; TITLE OF INVENTION: OSTEOPROTEGERIN
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,022
; FILING DATE: 12-DEC-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/577,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-378
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 401 amino acids
; TYPE: amino acid

Query Match 55.9%; Score 142; DB 3; Length 401;
Best Local Similarity 52.3%; Pred. No. 1.4e-07;
Matches 23; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 1 CPPGAGVIAGTPTSQNTQCPCPPGTFSASSSSSEQCQPHRNCT 44
|||||
Db 124 CPPGLGLVQAGTPTERTVCKPCDGFSGTSSKAPCKKHTNCS 167
|||||

RESULT 3
US-08-974-022-4
; Sequence 4, Application US/08974022
; Patent No. 6015938
; GENERAL INFORMATION:
; APPLICANT: Boyle, William J.
; APPLICANT: Lacey, David L.
; APPLICANT: Calzone, Frank J.
; APPLICANT: Chang, Ming-Shi
; TITLE OF INVENTION: OSTEOPROTEGERIN
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks


```

US-08-385-229-2
Query Match          53.5%; Score 136; DB 1; Length 461;
Best Local Similarity 51.2%; Pred. No. 6.8e-07;
Matches 22; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

Qy 1 CPPGAGVIAPGTQSNTQCPCPGTGSASSSSSEQQCPHRNC 43
Db 143 CRPFGVARGPTETSDVVCKPCAFPTSNTTTSSTDICRPHQIC 185
      ||| || |:: : ||| |||| ::||: ||:: ||:: |
      143 CRPFGVARGPTETSDVVCKPCAFPTSNTTTSSTDICRPHQIC 185

RESULT 9
US-08-650-000-2
; Sequence 2, Application US/08650000
; Patent No. 5945397
; GENERAL INFORMATION:
; APPLICANT: Smith, Craig A.
; APPLICANT: Goodwin, Raymond G.
; APPLICANT: Beckmann, M. Patricia
; TITLE OF INVENTION: Tumor Necrosis Factor Receptors
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/650,000
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,453
; FILING DATE:
; APPLICATION NUMBER: US/08/038,765
; FILING DATE:
; APPLICATION NUMBER: US 403,241
; FILING DATE: 05-SEP-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 405,370
; FILING DATE: 11-SEP-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 421,417
; FILING DATE: 13-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 523,635
; FILING DATE: 10-MAY-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Wight, Christopher L.
; REGISTRATION NUMBER: 31,680
; REFERENCE/DOCKET NUMBER: 2501-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 461 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-650-000-2

Query Match          53.5%; Score 136; DB 2; Length 461;
Best Local Similarity 51.2%; Pred. No. 6.8e-07;
Matches 22; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

```

QY 1 CPFGAGVIAPGTPSNTQCPCPGTFSASSSSSQCPHNC 43
 - - - - - : - - - - - : - - - - - : - - - - -
D6 143 CRPFGVARGPTETSDVVKPCAPGTSTTTSTDICRPHOIC 185

```

RESULT 10
5395760-2
; Patent NO. 5395760
; APPLICANT: SMITH, CRAIG A.; GOODWIN, RAYMOND G.; BECKMANN,
; M. PATRICIA
; TITLE OF INVENTION: DNA ENCODING TUMOR NECROSIS FACTOR-a AND
; B-RECEPTORS
; NUMBER OF SEQUENCES: 17
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/523,635
; FILING DATE: 10-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 421,417
; FILING DATE: 13-OCT-1989
; APPLICATION NUMBER: 405,370
; FILING DATE: 11-SEP-1989
; APPLICATION NUMBER: 403,241
; FILING DATE: 05-SEP-1989
; SEQ ID NO: 2
; LENGTH: 461
5395760-2

```

```

Query Match      53.5%;   Score 136;   DB 5;   Length 461;
Best local Similarity 51.2%;   Pred. No. 6.8e-07;
Matches 22;   Conservative 9;   Mismatches 12;   Indels

QY      1  CPPGAGVIAPCTSQNTQCPCPPGCTFSASSSSSQCCPHRNC 43
      | | | | | | | | | | | | | | | | | | | |
Db      143 CRPGFGVARPCTSDVYVKCPACPGCTFNSTTSDICRPHQIC 185

```

RESULT 11
US-08-243-010-1
; Sequence 1, Application US/08243010
; Patent No. 5639597
; GENERAL INFORMATION:
; APPLICANT: Laufer, Leander
; APPLICANT: Zetlimeissel, Gerd
; APPLICANT: Quendo, Patricia
; TITLE OF INVENTION: Cell-free Receptor Binding Assays, The
; TITLE OF INVENTION: Production and Use Thereof
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/243.010
; FILING DATE: 13-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/798,564
; FILING DATE: 26-NOV-1991
; APPLICATION NUMBER: DE P 40 37 837.3
; FILING DATE: 28-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Einaudi, Carol P.

```

;      REGISTRATION NUMBER: 32,220
;      REFERENCE/DOCKET NUMBER: 02481-1132-00000
;      TELECOMMUNICATION INFORMATION:
;      TELEPHONE: 202-408-4000
;      TELEFAX: 202-408-4400
;      INFORMATION FOR SEQ ID NO: 1:
;      SEQUENCE CHARACTERISTICS:
;      LENGTH: 486 amino acids
;      TYPE: amino acid
;      TOPOLOGY: linear
;      MOLECULE TYPE: peptide
;      US-08-243-010-1

```

```

Query Match      53.5%; Score 136; DB 1; Length 486;
Best Local Similarity 51.2%; Pred. No. 7.2e-07;
Matches 22; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

Qy 1 CPPGAGVIAPCTSONTCQCPGCTFSASSSSSEQCQPHRNC 43
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 143 CRPGFGVAPGCTSDVVVVCAPGCTGNTTSSDTCRPHQIC 185
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

RESULT 12
US-08-385-229-4
; Sequence 4, Application US/08385229
; Patent No. 5605690
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Cindy A.
; APPLICANT: Smith, Craig A.
; TITLE OF INVENTION: Method of Treating TNF-Dependent
; TITLE OF INVENTION: Inflammation Using Tumor Necrosis Factor Antagonists
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/385,229
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/946,236
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wight, Christopher L.
; REGISTRATION NUMBER: 31,680
; REFERENCE/DOCKET NUMBER: 2503
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 587-0606
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 518 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-385-229-4

Query Match 53.5%; Score 136; DB 1; Length 518;
Best Local Similarity 51.2%; Pred. No. 7.7e-07;
Matches 22; Conservative 9; Mismatches 12; Indels 0; Gaps 0;
Qv 1 CPPGAGVIAGTPTSONTCOPCPCTFSASSSSSEOCOPHRC 43

RESULT 14
5395760-4

Query Match 42.1%; Score 107; DB 3; Length 197;
Best Local Similarity 46.5%; Pred. No. 0.00026;
Matches 20; Conservative 5; Mismatches 18; Indels 0; Gaps 0;

QY 1 CPPGAGVIAPGTPSONTOCQPCPPGTFSSASSSSSEQCQPHRNC 43
DB 125 CIPGFGVMEMATETTTDTVCHPCVPVGFFSNQSSLFKCYPTWTS 167

Search completed: January 30, 2001, 16:47:42
Job time: 618 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:45:22 ; Search time 183.67 Seconds
(without alignments)
1.862 Million cell updates/sec

Title: US-09-518-931-2_COPY_185_194

Perfect score: 55

Sequence: 1 GSSSHDTLCT 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_36:*

1:	/SIDS1/gcgdata/geneseq/geneseq/AA1980.DAT:*
2:	/SIDS1/gcgdata/geneseq/geneseq/AA1981.DAT:*
3:	/SIDS1/gcgdata/geneseq/geneseq/AA1982.DAT:*
4:	/SIDS1/gcgdata/geneseq/geneseq/AA1983.DAT:*
5:	/SIDS1/gcgdata/geneseq/geneseq/AA1984.DAT:*
6:	/SIDS1/gcgdata/geneseq/geneseq/AA1985.DAT:*
7:	/SIDS1/gcgdata/geneseq/geneseq/AA1986.DAT:*
8:	/SIDS1/gcgdata/geneseq/geneseq/AA1987.DAT:*
9:	/SIDS1/gcgdata/geneseq/geneseq/AA1988.DAT:*
10:	/SIDS1/gcgdata/geneseq/geneseq/AA1989.DAT:*
11:	/SIDS1/gcgdata/geneseq/geneseq/AA1990.DAT:*
12:	/SIDS1/gcgdata/geneseq/geneseq/AA1991.DAT:*
13:	/SIDS1/gcgdata/geneseq/geneseq/AA1992.DAT:*
14:	/SIDS1/gcgdata/geneseq/geneseq/AA1993.DAT:*
15:	/SIDS1/gcgdata/geneseq/geneseq/AA1994.DAT:*
16:	/SIDS1/gcgdata/geneseq/geneseq/AA1995.DAT:*
17:	/SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT:*
18:	/SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT:*
19:	/SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT:*
20:	/SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT:*
21:	/SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	100.0	153	20	Human TNFR superfa
2	55	100.0	215	20	Human HAP06 protei
3	55	100.0	245	20	A human tumour nec
4	55	100.0	271	20	Human mFLINT #1 pr
5	55	100.0	273	20	Human mFLINT #2 pr
6	55	100.0	300	19	Amino acid sequenc
7	55	100.0	300	19	Human tumour necro
8	55	100.0	300	20	Human lung TNF-rec
9	55	100.0	300	20	Human FLINT #1 pro
10	55	100.0	300	20	Mammalian tumour n
11	55	100.0	300	20	Human Dcr3 polyep
12	55	100.0	300	20	Human tumour necro

13	55	100.0	300	20	Orphan receptor (H
14	55	100.0	300	21	Human TNF receptor
15	55	100.0	302	21	Human FLINT #2 pro
16	37	67.3	33	18	MAB anti-HBsAg bin
17	37	67.3	225	20	Tobacco RTG1 homol
18	37	67.3	467	21	AV37 antigen prote
19	36	65.5	75	20	TNF-R extracellular
20	36	65.5	140	20	Polypeptide fragme
21	36	65.5	223	20	Tobacco RTG1 homol
22	36	65.5	276	20	Human CD40 protein
23	36	65.5	277	14	CD40 protein. Hom
24	36	65.5	277	20	Human CD40 protein
25	36	65.5	277	20	Human CD40 protein
26	36	65.5	364	20	Polypeptide fragme
27	36	65.5	1245	13	Bacillus thuringie
28	36	65.5	1245	18	Bacillus thuringie
29	36	65.5	1245	18	B.t. toxin protein
30	36	65.5	1245	19	86Q3a toxin. Baci
31	35	63.6	134	17	Human TNF receptor
32	35	63.6	136	19	Human TNF receptor
33	35	63.6	186	21	Soluble herpesviro
34	35	63.6	187	17	Mutated OCIF, OCIF
35	35	63.6	187	21	Primate protein se
36	35	63.6	193	21	Soluble herpesviro
37	35	63.6	197	17	Mutated OCIF, OCIF
38	35	63.6	197	21	Soluble herpesviro
39	35	63.6	272	17	Mutated OCIF, OCIF
40	35	63.6	277	21	Membrane-bound her
41	35	63.6	283	17	Human tumour necro
42	35	63.6	283	18	Human herpes simpl
43	35	63.6	283	19	Herpesvirus entry
44	35	63.6	283	19	Human TNF receptor
45	35	63.6	283	20	Human tumour-assoc

ALIGNMENTS

RESULT 1

Y22222

ID Y22222 standard; Protein; 153 AA.

AC Y22222;

DF 16-SEP-1999 (first entry)

DE Human TNFR superfamily soluble receptor protein sequence.

KW TNFL1; human; TNFR superfamily; tumour necrosis factor ligand; TNF;

KW tumour necrosis factor receptor; TNFR superfamily; cell proliferation;

KW cell differentiation; cytokine production; immunoglobulin; hyperplasia;

KW apoptosis inducer; activated T cell; autoimmune disease; inhibitor;

KW rheumatoid arthritis; insulin-dependent diabetes mellitus; endotoxin shock;

KW lymphoid organogenesis; bacterial resistance; contact hypersensitivity;

KW delayed type sensitivity; therapy.

OS Homo sapiens.

PN WO9933980-A2.

PD 08-JUL-1999.

PF 22-DEC-1998; 98WO-US27474.

PR 16-DEC-1998; 98US-0212270.

PR 30-DEC-1997; 97US-0068959.

PA (CHIR) CHIRON CORP.

PI Kassam A, Lamson G, Pot D, Tribouley C;

Wed Jan 31 13:43:58 2001

DR WPI: 1999-405508/34.
 DR N-PSDB; X84621.
 XX
 XX New tumour necrosis factor ligands, useful for induction of cell
 PT death and/or proliferation of cells
 PT
 XX Claim 1; Page 61; 69pp; English.
 PS
 XX This sequence represents a tumour necrosis factor receptor (TNFR)
 CC superfamily soluble protein of the invention. The invention also relates
 CC to tumour necrosis factor (TNF) ligand (TNFL) family proteins. The TNFL
 CC proteins play regulatory roles in cell proliferation and/or
 CC differentiation, e.g. they can induce production of cytokines,
 CC immunoglobulins, etc. A variety of diseases can be treated by modulating
 CC the activity of TNFL proteins, e.g. they can induce apoptosis of
 CC activated T cells but rescue resting T cell from apoptosis. TNFL
 CC polypeptides can therefore be used to treat autoimmune diseases, such as
 CC myasthenia gravis, insulin-dependent diabetes mellitus, rheumatoid
 CC arthritis, multiple sclerosis, and systemic lupus erythematosus. TNFL
 CC proteins also have tumour stimulating properties, so tumours can be
 CC treated by inhibiting the expression or activity of TNFL. Other
 CC proliferative disorders, such as neoplasias, dysplasias, and hyperplasia
 CC can also be treated using TNFL inhibitors. The TNFL polypeptides and
 CC polynucleotides can also be used to enhance or decrease TNF activity,
 CC thus providing therapeutic benefits such as induction of cell death,
 CC lymphoid organogenesis, or host bacterial resistance, and inhibition of
 CC endotoxic shock, contact hypersensitivity, delayed type sensitivity or
 CC immunocompetence of a transplant recipient. TNF and its receptors play a
 CC major role in host defence and immunosurveillance. As such, there is a
 CC need to identify new members of TNFR families. This invention provides
 CC this need.
 XX Sequence 153 AA:
 SQ

Query Match 100.0%; Score 55; DB 20; Length 153;
 Best Local Similarity 100.0%; Pred. No. 0.0043;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSSHDTLCT 10
 |||||
 DB 101 gssshdtlct 110

RESULT 2
 W93585
 ID W93585 standard; Protein; 215 AA.
 AC W93585;
 XX
 XX 18-JUN-1999 (first entry)
 DT
 XX Human hAPO6 protein.
 DE
 XX Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
 KW developmental abnormality; gestational abnormality; prostate cancer;
 KW APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
 KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
 KW apoptosis; human.
 XX
 XX Homo sapiens.
 OS
 XX W09911791-A2.
 PN
 XX 11-MAR-1999.
 PD
 XX .04-SEP-1998; 98WO-US18393.
 PF
 XX 05-SEP-1997; 97US-0924634.
 PR
 XX (UNIV) UNIV WASHINGTON.
 PA
 XX Chaudhary PM;
 PI

XX WPI: 1999-205191/17.
 DR N-PSDB; X23419.
 XX
 XX New Tumor Necrosis Factor family receptor polypeptides and ligands -
 PT useful for diagnosis and treatment of prostate cancer and
 PT developmental or gestational abnormalities
 PT
 XX Claim 29; Fig 9; 156pp; English.
 PS
 XX This invention describes isolated Tumor Necrosis Factor (TNF) family
 CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
 CC their active fragments. APO4 is useful for diagnosing prostate cancer
 CC by determining levels of APO4 in an individual. Prostate cancer can also
 CC be treated using APO4 selective binding agents linked to a therapeutic
 CC moiety. APO4 polypeptides are also useful for identifying selective
 CC binding agents, useful in diagnosis/treatment of disease by binding of
 CC agents to the polypeptide/active fragment which is extracellular, or
 CC expressed on the cell surface. The binding is preferably performed in
 CC vivo. APO4 polypeptides/ active fragments are also useful for screening
 CC for agonists and antagonists by binding and observing the change in APO4
 CC activity. Effective pharmacological agents useful in diagnosis or
 CC treatment of disease are also identified using APO4 polypeptides/active
 CC fragments and APO4 signal transducer molecules that specifically interact
 CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4
 CC activity. The method is performed in vivo or in vitro. APO polypeptides
 CC are all useful as immunogens for preparing antibodies. APO4 is also
 CC useful for diagnosis/treatment of developmental or gestational
 CC abnormalities. APO8 was transfectected to human breast carcinoma cell line
 CC MCF-7, and induced apoptosis.
 XX Sequence 215 AA:
 SQ

Query Match 100.0%; Score 55; DB 20; Length 215;
 Best Local Similarity 100.0%; Pred. No. 0.0061;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSSHDTLCT 10
 |||||
 DB 100 gssshdtlct 109

RESULT 3
 Y28449
 ID Y28449 standard; Protein; 245 AA.
 AC Y28449;
 XX
 XX 29-SEP-1999 (first entry)
 DT
 XX A human tumour necrosis factor-R2-like proteins (TR2P)-1.
 DE
 XX Human tumour necrosis factor-R2-like protein; TR2P; achondroplasia;
 KW osteoporosis; developmental disorder; Cushing's syndrome;
 KW muscular dystrophy; epilepsy; hereditary neuropathy;
 KW Charcot-Marie-Tooth disease; neurofibromatosis; hypothyroidism;
 KW hydrocephalus; seizure disorder; cerebral palsy; spinal bifida;
 KW congenital glaucoma; cataract; sensorineural hearing loss;
 KW reproductive disorder; infertility; ovulatory defect; endometriosis;
 KW autoimmune disorder; ectopic pregnancy; teratogenesis; spermatogenesis;
 KW immunological disorder; AIDS; Addison's disease; allergy; bronchitis;
 KW atherosclerosis; diabetes mellitus; Chron's disease; lupus;
 KW irritable bowel syndrome; multiple sclerosis; infection;
 KW neoplastic disorder; adenocarcinoma; leukaemia; lymphoma; melanoma;
 KW myeloma; sarcoma.
 XX
 XX Homo sapiens.
 OS
 XX W09931128-A2.
 PN
 XX 24-JUN-1999.
 XX PD


```

XX 02-DEC-1998; 98WO-US25649.
XX
XX
XX 16-DEC-1997; 97US-0991945.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Au-Young J, Bandman O, Hillman JL, Kaser MR, Tang YT;
XX
XX WPI; 1999-457916/38.
XX
XX N-PSDB; X89503.
XX
XX New tumour necrosis factor-R2-like protein - useful in the treatment
XX of osteogenesis, developmental, reproductive, immunological and
XX neoplastic disorders
XX
XX Claim 1; Fig 1A-C; 8lpp; English.
XX
XX The present sequence represents a human tumour necrosis factor-R2-like
XX protein (TR2P). The protein is used to treat and prevent osteogenesis,
XX developmental, reproductive, immunological and neoplastic disorders, and
XX also to diagnose disorders associated with TR2 protein expression. Such
XX disorders include osteogenesis disorders such as achondroplasia and
XX osteoporosis, developmental disorders such as Cushing's syndrome,
XX muscular dystrophy, and epilepsy, hereditary neuropathies such as
XX Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism,
XX hydrocephalus, seizure disorders such as cerebral palsy and spinal
XX bifida, congenital glaucoma, cataract, or sensorineural hearing loss,
XX reproductive disorders such as infertility, ovulatory defects and
XX endometriosis, autoimmune disorders, ectopic pregnancy and teratogenesis,
XX disruption of spermatogenesis, immunological disorders such as AIDS,
XX Addison's disease, allergies, bronchitis, atherosclerosis, diabetes
XX mellitus, Chron's disease, lupus and irritable bowel syndrome, multiple
XX sclerosis, viral, fungal, helminthic, parasitic and protozoal infections,
XX and neoplastic disorders including adenocarcinoma, leukaemia, lymphoma,
XX melanoma, myeloma, sarcoma, and teratocarcinoma.
XX
XX Sequence 245 AA;
XX
XX Query Match 100.0%; Score 55; DB 20; Length 245;
XX Best Local Similarity 100.0%; Pred. No. 0.007;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GSSSHDTLCT 10
XX Db 185 gssshdtlct 194
XX
XX RESULT 4
XX Y42184
XX ID Y42184 standard; Protein; 271 AA.
XX
XX AC Y42184;
XX
XX DT 17-DEC-1999 (first entry)
XX
XX DE Human mFLINT #1 protein sequence.
XX
XX Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
XX apoptosis; inflammation; cancer; diabetes; acute liver failure;
XX sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
XX reperfusion-associated injury; aplastic anaemia; differentiation;
XX growth; myelodysplastic syndrome; pancytopenic condition;
XX myocardial ischaemia.
XX
XX OS Homo sapiens.
XX
XX PN WO950413-A2.
XX
XX PD 07-OCT-1999.
XX
XX PF 30-MAR-1999; 99WO-US06797.

```

```

XX 30-MAR-1998; 98US-0079856.
XX
XX 20-MAY-1998; 98US-0086074.
XX
XX 09-SEP-1998; 98US-0099643.
XX
XX 17-DEC-1998; 98US-0112577.
XX
XX 18-DEC-1998; 98US-0112703.
XX
XX 18-DEC-1998; 98US-0112933.
XX
XX 22-DEC-1998; 98US-0113407.
XX
XX (ELIL ) LILLY & CO ELI.
XX
XX Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
XX Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
XX Song HY, Wang J, Wu X, Zuckerman SH;
XX
XX WPI; 1999-591319/50.
XX
XX N-PSDB; Z25377.
XX
XX Use of mature FLINT for treating acute liver failure, inflammation,
XX cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
XX and proinflammatory activity
XX
XX Claim 31; Fig 3; 99pp; English.
XX
XX The present invention describes therapeutic applications of mature FLINT
XX (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
XX (mFLINT), which is a member of the tumour necrosis factor receptor
XX superfamily, is used for treating acute liver failure, inflammation of
XX the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
XX with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
XX injury or disorder such as hypercoagulation (including use with
XX thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
XX or disorder, Type I diabetes, cancer, cell damage or damage to an
XX innocent bystander tissue that is induced by a chemotherapeutic agent or
XX therapeutic irradiation, treating haematopoietic progenitor cells that
XX have been exposed to therapeutic radiation or chemotherapy, aplastic
XX anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
XX also used for promoting the growth or differentiation of a haematopoietic
XX progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
XX resulting from abnormal myocardial ischaemia. The present sequence
XX represents human mFLINT.
XX
XX Sequence 271 AA;
XX
XX Query Match 100.0%; Score 55; DB 20; Length 271;
XX Best Local Similarity 100.0%; Pred. No. 0.0077;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GSSSHDTLCT 10
XX Db 156 gssshdtlct 165
XX
XX RESULT 5
XX Y42185
XX ID Y42185 standard; Protein; 273 AA.
XX
XX AC Y42185;
XX
XX DT 17-DEC-1999 (first entry)
XX
XX DE Human mFLINT #2 protein sequence.
XX
XX Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
XX apoptosis; inflammation; cancer; diabetes; acute liver failure;
XX sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
XX reperfusion-associated injury; aplastic anaemia; differentiation;
XX growth; myelodysplastic syndrome; pancytopenic condition;
XX myocardial ischaemia.
XX
XX OS Homo sapiens.
XX
XX

```

PN WO950413-A2.
XX 07-OCT-1999.
XX 30-MAR-1999; 99WO-US06797.
XX 30-MAR-1998; 98US-0079856.
XX 20-MAY-1998; 98US-0086074.
XX 09-SEP-1998; 98US-0099643.
XX 17-DEC-1998; 98US-0112577.
XX 18-DEC-1998; 98US-0112703.
XX 18-DEC-1998; 98US-0112933.
XX 22-DEC-1998; 98US-0113407.
XX (ELIL) LILLY & CO ELI.
XX Bumol JF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
PI Hul KY, Kharitonovkov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
PI Song HY, Wang X, Wu X, Zuckerman SH;
XX WPI: 1999-591319/50.
XX N-PSDB: 225378.
XX Use of mature FLINT for treating acute liver failure, inflammation,
PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
PT and proinflammatory activity
XX Example 2: Fig 4; 99pp; English.
XX The present invention describes therapeutic applications of mature FLINT
CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
CC (mFLINT), which is a member of the tumour necrosis factor receptor
CC superfamily, is used for treating acute liver failure, inflammation of
CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
CC injury or disorder such as hypercoagulation (including use with
CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
CC or disorder. Type I diabetes, cancer, cell damage or damage to an
CC innocent bystander tissue that is induced by a chemotherapeutic agent or
CC therapeutic irradiation, treating haematopoietic progenitor cells that
CC have been exposed to therapeutic radiation or chemotherapy, aplastic
CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
CC also used for promoting the growth or differentiation of a haematopoietic
CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
CC resulting from abnormal myocardial ischaemia. The present sequence
CC represents human mFLINT.
XX
SQ Sequence 273 AA;

Query Match 100.0%; Score 55; DB 20; Length 273;
Best Local Similarity 100.0%; Pred. No. 0.0078;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSSHDTLCT 10
Db 158 gssshdtlct 167
|||||

RESULT 6
W66102
ID W66102 standard; Protein: 300 AA.
XX
XX AC W66102;
XX
XX 02-DEC-1998 (first entry)
DE Amino acid sequence of tumour necrosis related receptor (TR4).
XX Human; tumour necrosis related receptor; TR4; agonist; antagonist;
KW inhibition; chronic; acute; inflammation; arthritis; septicaemia;
KW autoimmune disease; transplant rejection; stroke; cancer;
KW Alzheimer's disease.

XX OS Homo sapiens.
XX EP861850-A1.
XX 02-SEP-1998.
XX 20-JAN-1998; 98EP-0300382.
XX 04-FEB-1997; 97US-0794796.
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX Emery J, Tan KB, Truneh A, Young PR;
XX WPI: 1998-508248/44.
XX N-PSDB: V07654.
XX New DNA encoding tumour necrosis related receptor - used to treat
PT and prevent e.g. inflammation, arthritis, septicaemia, autoimmune
PT diseases, transplant rejection, infection, stroke, ischaemia, ARDS,
PT restenosis, AIDS, bone disorders and cancer
XX Claim 1; Fig 1; 2lpp; English.
XX This is the amino acid sequence of the human tumour necrosis related
CC receptor (TR4), used in the method of the invention. The TR4 protein
CC or its agonist can be used to treat a subject in need of enhanced
CC TR4 polypeptide activity. The antagonist is used to inhibit TR4
CC polypeptide activity. The active agents can be used for the
CC treatment and prevention of diseases such as chronic and acute
CC inflammation, arthritis, septicaemia, autoimmune diseases, transplant
CC rejection, stroke, cancer, Alzheimer's disease.
XX Sequence 300 AA;

Query Match 100.0%; Score 55; DB 19; Length 300;
Best Local Similarity 100.0%; Pred. No. 0.0086;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSSHDTLCT 10
Db 185 gssshdtlct 194
|||||

RESULT 7
W63622
ID W63622 standard; Protein: 300 AA.
XX
XX AC W63622;
XX
XX 26-OCT-1998 (first entry)
DE Human tumour necrosis factor receptor-6 alpha protein.
XX Human tumour necrosis factor receptor-6 alpha; TNFR-6 alpha; TNFR-6 beta;
KW endothelial cells; keratinocytes; normal prostate; apoptosis;
KW prostate tumour tissue.
XX
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..30
XX Protein 31..300
XX /note= "TNFR-6 alpha"
XX Region 31..282
XX /note= "Soluble extracellular domain"
XX WO9830694-A2.
XX 16-JUL-1998.
XX

```

PF 13-JAN-1998; 98WO-US00153.
XX
PR 14-JAN-1997; 97US-0035496.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ebner R, Feng P, Gentz RL, Ni J, Ruben SM, Yu G;
XX
DR WPI; 1998-399142/34.
DR N-PSDB; V39085.
XX
PT Human tumour necrosis factor receptors 6-alpha and 6-beta - used in
PT the diagnosis of immune system-related disorder(s)
XX
PS Claim 20; Fig 1; 9lpp; English.
XX
CC The present sequence represents the human tumour necrosis factor
CC receptor-6 alpha (TNFR-6 alpha) protein. The invention also provides
CC for the TNFR-6 beta protein (W63623). TNFR-6 alpha and TNFR-6 beta are
CC members of the tumour necrosis factor receptor (TNFR) family. TNFRs
CC are expressed in endothelial cells, keratinocytes, normal prostate and
CC prostate tumour tissue. For a number of disorders of these cells,
CC particularly of the immune system, substantially altered (whether
CC increased or decreased) levels of TNFR-6 alpha and/or TNFR-6 beta gene
CC expression can be detected, therefore the TNFR-6 alpha and TNFR-6 beta
CC polypeptides, nucleic acids and antibodies are claimed to be useful in
CC the diagnosis of such disorders. Mutations of the TNFR-6 alpha and
CC TNFR-6 beta genes can also be detected. The TNFR polypeptides are
CC also claimed to be useful for identifying ligands which may be useful
CC in the treatment of apoptosis related disorders.
XX
SQ Sequence 300 AA;

Query Match 100.0%; Score 55; DB 19; Length 300;
Best Local Similarity 100.0%; Pred. No. 0.0086;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSSHDTLCT 10
Db 185 gssshdtlct 194
|||||

RESULT 8
ID Y03099 standard; Protein; 300 AA.
XX
AC Y03099;
XX
DT 09-DEC-1999 (first entry)
XX
DE Human lung TNF-receptor protein.
XX
KW Tumour necrosis factor; TNF; TNF receptor; human; lung; gene therapy;
KW detection; immunoassay; diagnosis; disease; immune system; tumour;
KW osteogenic system; cardiovascular system; central nervous system; asthma;
KW peripheral nervous systems; transplant incompatibility; antitumor;
KW Rheumatoid arthritis; antiasthmatic; antiarthritic.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 134..1036
FT /*tag= a
FT /product= "TNF-receptor"
XX
XX DE19809978-A1.
XX
PN 16-SEP-1999.
XX
PD
XX
PF 09-MAR-1998; 98DE-1009978.
XX
XX 09-MAR-1998; 98DE-1009978.
XX
PR

```

```

XX (BADI ) BASF AG.
XX
XX Kroeger B;
XX
DR WPI; 1999-519473/44.
DR N-PSDB; Z09998.
XX
PT New soluble member of tumor necrosis factor receptor family, useful for
PT identification specific modulators and for treating disease e.g. tumors
XX
XX Claim 1; Page 8-9; 10pp; German.
XX
CC This invention describes a novel tumour necrosis factor (TNF) receptor
CC (I) isolated from human lung tissue. (I) is used: (i) to raise specific
CC antibodies (Ab); (ii) to screen for specific (antagonists or ligands
CC (A), potential therapeutic agents; and (iii) therapeutically (optionally
CC expressed from a gene therapy vector) in conditions associated with a
CC deficit of (I). Ab are used: (a) for qualitative or quantitative
CC detection of (I) in standard immunoassays (for diagnosis of disease, or
CC susceptibility, or for monitoring); and (b) as therapeutic inhibitors in
CC cases where (I) is overexpressed. Nucleic acid (II) that encodes (I) is
CC used: (A) for recombinant production of (I); (B) also its oligonucleotide
CC fragments, in standard hybridization and/or amplification assays; (C) as
CC source of antisense molecules or ribozymes; and (D) to produce transgenic
CC animals (for studying (patho)physiology of (I)). Diseases possibly
CC associated with under- or over-expression of (I) are those of the immune,
CC osteogenic, cardiovascular and central or peripheral nervous systems,
CC tumors, transplant incompatibility, asthma and rheumatoid arthritis. The
CC products of the invention have antitumor, antiasthmatic and
CC antiarthritic activity. This sequence represents the TNF-receptor of the
XX
XX invention.
XX
SQ Sequence 300 AA;

Query Match 100.0%; Score 55; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 0.0086;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSSHDTLCT 10
Db 185 gssshdtlct 194
|||||

RESULT 9
ID Y42182 standard; Protein; 300 AA.
XX
AC Y42182;
XX
DT 17-DEC-1999 (first entry)
XX
DE Human FLINT #1 protein sequence.
XX
KW Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
KW reperfusion-associated injury; aplastic anaemia; differentiation;
KW growth; myelodysplastic syndrome; pancytopenic condition;
KW myocardial ischaemia.
XX
OS Homo sapiens.
XX
PN WO9950413-A2.
XX
PD 07-OCT-1999.
XX
PF 30-MAR-1999; 99WO-US06797.
XX
XX 30-MAR-1998; 98US-0079856.
XX
PR 20-MAY-1998; 98US-0086074.

```

```
PR 09-SEP-1998; 98US-0099643.
PR 17-DEC-1998; 98US-0112577.
PR 18-DEC-1998; 98US-0112703.
PR 18-DEC-1998; 98US-0112933.
PR 22-DEC-1998; 98US-0113407.
XX (ELIL ) LILLY & CO ELI.
PA Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
PI Song HY, Wang J, Wu X, Zuckerman SH;
XX WPI: 1999-591319/50.
DR N-PSDB; Z25375.
XX
XX Use of mature FLINT for treating acute liver failure, inflammation,
PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
PT and proinflammatory activity
XX
XX Claim 30; Fig 1: 99pp; English..
XX
XX The present invention describes therapeutic applications of mature FLINT
CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
CC (mFLINT), which is a member of the tumour necrosis factor receptor
CC superfamily, is used for treating acute liver failure, inflammation of
CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
CC injury or disorder such as hypercoagulation (including use with
CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
CC or disorder, Type I diabetes, cancer, cell damage or damage to an
CC innocent bystander tissue that is induced by a chemotherapeutic agent or
CC therapeutic irradiation, treating haematopoietic progenitor cells that
CC have been exposed to therapeutic radiation or chemotherapy, aplastic
CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
CC also used for promoting the growth or differentiation of a haematopoietic
CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
CC resulting from abnormal myocardial ischaemia. The present sequence
CC represents human FLINT.
XX
XX Sequence 300 AA;
SQ
Query Match 100.0%; Score 55; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 0.0086;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GSSSHDTLCT 10
Db 185 gssshdtlct 194
RESULT 10
Y17479
ID Y17479 standard; Protein; 300 AA.
XX
XX Y17479;
AC
XX
XX 02-AUG-1999 (first entry)
DT
XX
XX Mammalian tumour necrosis factor receptor OPG-2.
DE
XX
XX Tumour necrosis factor receptor; TNF receptor; OPG-2; Paget's disease;
KW osteopenic disorder; osteoclast activity; primary osteoporosis;
KW hyperglycaemia; osteolytic metastasis; immune response; cancer.
XX
XX Mammalia.
OS
XX
XX WO9926977-A1.
PN
XX
XX 03-JUN-1999.
PD
XX
XX 24-NOV-1998; 98WO-US25065.
PF
XX
```

```
PR 17-FEB-1998; 98US-0074896.
PR 24-NOV-1997; 97US-0066446.
XX
XX (BIOJ ) BIOGEN INC.
PA
XX
XX Tschopp J;
PI
XX
XX WPI: 1999-347693/29.
DR N-PSDB; X76052.
XX
XX New tumour necrosis factor family receptor OPG-2
PT
XX
XX Claim 1; Page 18; 22pp; English.
XX
XX The present sequence represents a mammalian tumour necrosis factor
CC receptor, designated OPG-2. OPG-2, is a member of the tumour necrosis
CC factor receptor family, and can be used: (i) to raise specific
CC antibodies (Ab), (ii) to treat osteopenic disorders associated with
CC excessive osteoclast activity, e.g. primary osteoporosis, Paget's
CC disease, hyperglycaemia of malignancy, or osteolytic metastases; (iii)
CC for affinity purification of cognate ligands, and (iv) to screen for
CC ligands (antagonists or agonists). Ab, or other OPG-2 blocking agents
CC such as soluble forms of the protein, are used to prevent, or reduce
CC severity of, an immune response, and for treating cancer. They can also
CC be used in diagnostic assays. The nucleic acid sequence encoding OPG-2
CC can be used as a probe to isolate related sequences from other species.
XX
XX Sequence 300 AA;
SQ
Query Match 100.0%; Score 55; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 0.0086;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GSSSHDTLCT 10
Db 185 gssshdtlct 194
RESULT 11
Y06817
ID Y06817 standard; Protein; 300 AA.
XX
XX Y06817;
AC
XX
XX 24-JUN-1999 (first entry)
DT
XX
XX Human DcR3 polypeptide.
DE
XX
XX DcR3 polypeptide; tumour necrosis factor receptor; TNFR; Fas ligand;
KW apoptosis; T cell mediated immune response; allergy; asthma; cancer;
KW rheumatoid arthritis; Crohn's disease; guest vs. host disease; human;
KW gene therapy.
XX
XX Homo sapiens.
OS
XX
XX WO9914330-A1.
PN
XX
XX 25-MAR-1999.
PD
XX
XX 18-SEP-1998; 98WO-US19661.
PF
XX
XX 30-JUL-1998; 98US-0094640.
PR
XX
XX 18-SEP-1997; 97US-0059288.
PR
XX
XX (GETH ) GENENTECH INC.
PA
XX
XX Ashkenazi AJ, Botstein D, Dodge KH, Goddard A, Gurney AL;
PI Kim KJ, Lawrence DA, Pitti R, Roy MA, Tumas DB;
PI Wood WI;
XX
XX WPI: 1999-244032/20.
DR
XX
XX N-PSDB; X32744.
```

XX Dcr3 polypeptide related to tumor necrosis factor receptor
XX
XX
PS Claim 5; Fig 1; 88pp; English.
XX
CC This represents a human Dcr3 polypeptide, a homologue of tumour necrosis
CC factor receptor (TNFR) polypeptide. Host cells containing a vector
CC comprising the Dcr3 nucleic acid can be used for the recombinant
CC expression of the protein. Dcr3 binds to Fas ligand, so it (or its
CC chimeras) are useful for modulating apoptosis in mammalian cells, also
CC other Fas-ligand induced activities, particularly to inhibit T cell
CC mediated immune responses, e.g. in treatment of allergy, asthma,
CC rheumatoid arthritis, Crohn's disease, guest vs. host disease etc. Dcr3
CC may also be used to identify specific binding proteins, potential
CC inhibitors. Antibodies against Dcr3 are used to treat cancer,
CC specifically of the lung and colon, also in diagnosis and for affinity
CC purification of the protein. Detecting mutations in the gene for Dcr3 is
CC also used to diagnose cancer, or predisposition to it. Dcr3 nucleic acid
CC is useful as hybridization probe to detect genomic or related sequences;
CC for chromosome and gene mapping; as source of antisense sequences; for
CC expression of recombinant Dcr3 and to generate transgenic animals (for
CC development and screening of therapeutic agents), also for in vivo or
CC ex vivo gene therapy.
XX
SQ Sequence 300 AA;

Query Match 100.0%; Score 55; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 0.0086;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSSHDTLCT 10
Db |||||||
Db 185 gssshdtlct 194

RESULT 12
W97749
ID W97749 standard; Protein; 300 AA.
XX
XX AC W97749;
XX
XX DT 21-MAY-1999 (first entry)
XX
XX DE Human tumour necrosis factor receptor ZTNFR-5.
XX
XX KW ZTNFR-5; tumour necrosis factor receptor; TNFR; human;
XX cell maturation; bone cell regulation.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX Peptide 1..23
XX FT /note= "signal peptide"
XX FT 24..300
XX FT /note= "mature protein"
XX FT 24..194
XX FT /note= "extracellular domain"
XX FT 49..71
XX FT /note= "cysteine-rich pseudo-repeat 1"
XX FT 72..113
XX FT /note= "cysteine-rich pseudo-repeat 1"
XX FT 114..151
XX FT /note= "cysteine-rich pseudo-repeat 1"
XX FT 152..194
XX FT /note= "cysteine-rich pseudo-repeat 1"
XX
XX PN WO9904001-A1.
XX
XX PD 28-JAN-1999.
XX
XX PF 21-JUL-1998; 98WO-US15072.

PR 21-JUL-1997; 97US-0053203.
XX
XX PA (ZYMO) ZYMOGENETICS INC.
XX
XX PI Farrah TM;
XX
XX DR WPI; 1999-132245/11.
XX N-PSDB; X07226.
XX
XX PT Novel tumour necrosis factor receptor ZTNFR5 - useful for
XX regulating maturation of TNF-ligand bearing cells
XX
XX PS Claim 1; Page 84-85; 109pp; English.
XX
XX CC This polypeptide comprises a new, secreted tumour necrosis factor
XX receptor (see W97749), designated ZTNFR-5. Novel ZTNFR-5 encoding
XX polynucleotides and polypeptides were initially identified by
XX querying an expressed sequence tag (EST) database for sequences
XX homologous to conserved motifs within the TNF receptor family.
XX Based on this search, a contig of 16 ESTs (see X07226) was
XX constructed. ZTNFR-5 polypeptides comprise 4 cysteine-rich repeats
XX (see also W97750-55) that are homologous to other TNF receptors, in
XX particular the soluble, secreted TNF receptor osteoprotegerin.
XX ZTNFR-5 polypeptide can be prepared by recombinant methods. The
XX polypeptide, especially the extracellular domain, can be used to
XX generate a soluble variant of ZTNFR-5. The polypeptides and
XX nucleic acids can be used to screen for ligands, agonists and
XX antagonists of ZTNFR-5. The polypeptides can be used in bone cell
XX regulation and to regulate the maturation of TNF ligand-bearing
XX cells such as T- or B-cells, lymphocytes, peripheral blood
XX mononuclear cells, polymorphonuclear leukocytes, fibroblasts or
XX haematopoietic cells.
XX
SQ Sequence 300 AA;

Query Match 100.0%; Score 55; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 0.0086;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSSHDTLCT 10
Db |||||||
Db 185 gssshdtlct 194

RESULT 13
W95082
ID W95082 standard; Protein; 300 AA.
XX
XX AC W95082;
XX
XX DT 20-MAY-1999 (first entry)
XX
XX DE Orphan receptor (HUMAN NTR-1) polypeptide.
XX
XX KW HUMAN NTR-1; orphan receptor; osteoprotegerin; OPG; TNFR; human;
XX tumour necrosis factor receptor; muscle disorder; bone mass; screening;
XX muscle metabolism; binding agent; cognate ligand.
XX
XX OS Homo sapiens.
XX
XX PN WO9907738-A2.
XX
XX PD 18-FEB-1999.
XX
XX PF 04-AUG-1998; 98WO-US16202.
XX
XX PR 06-AUG-1997; 97US-0054869.
XX
XX PA (PROC) PROCTER & GAMBLE CO.
XX (REGE-) REGENERON PHARM INC.
XX
XX PI Masiakowski PJ, Morris J, Valenzuela DM;

Wed Jan 31 13:43:58 2001

```

XX WPI; 1999-167365/14.
XX DR N-PSDB; X22300.
XX PT Novel orphan human receptor polypeptide and nucleic acid - useful as
XX PT diagnostic reagents and for treatment of muscle disorders
XX PS
XX PS Claim 7; Page 21; 23pp; English.
XX CC This represents a HUMAN NTR-1 polypeptide, a novel orphan receptor. The
XX CC protein is related to osteoprotegerin (OPG) and to tumour necrosis factor
XX CC receptor (TNFR). Host cells transformed with a vector comprising the
XX CC HUMAN NTR-1 nucleic acid are used for the recombinant expression of the
XX CC protein. HUMAN NTR-1 proteins and antibodies immuno specific for the
XX CC protein are useful for diagnosis and treatment of humans and animals,
XX CC especially muscle disorders, as the receptor is involved in regulation of
XX CC bone mass and muscle metabolism. HUMAN NTR-1 receptors are also useful
XX CC for screening for novel binding agents, and cognate ligands, which may be
XX CC used to treat disorders associated with HUMAN NTR-1 imbalance.
XX CC
XX CC Sequence 300 AA:
XX SQ
XX Query Match 100.0%; Score 55; DB 20; Length 300;
XX Best Local Similarity 100.0%; Pred. No. 0.0086;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GSSSHDTLCT 10
XX DB |||||||||
XX DB 185 gssshdtlct 194
XX
XX RESULT 14
XX ID Y77458 standard; Protein: 300 AA.
XX AC Y77458;
XX XX
XX DT 05-JUN-2000 (first entry)
XX DE Human TNF receptor-like protein, HDTEA84.
XX XX
XX KW TNF receptor family; tumour necrosis factor; HDTEA84; HSLJD37R;
XX KW Rank-like protein; RANKL; immune disorder; inflammation; allergy;
XX KW immunosuppressant; antiarthritic; antirheumatoid; antiinflammatory;
XX KW dermatological; antithyroid.
XX XX
XX OS Homo sapiens.
XX XX
XX PN WO200001817-A2.
XX XX
XX PD 13-JAN-2000.
XX XX
XX PF 06-JUL-1999; 99WO-US12366.
XX XX
XX PR 06-JUL-1998; 98US-0110938.
XX PR 13-JUL-1998; 98US-0114466.
XX PR 23-JUL-1998; 98US-0093897.
XX PR 12-AUG-1998; 98US-0132968.
XX PR 18-AUG-1998; 98US-0136214.
XX PR 11-SEP-1998; 98US-0099999.
XX XX
XX PA (SCHE ) SCHERING CORP.
XX XX
XX PI Bates EEM, Lebecque SJE, Murphy EE, Mattson JD, Gorman DM;
XX PI Hedrick JA, Wang L, Zlotnik A, Murgolo NJ, Greene JR, Johnston JA;
XX PI Bazan JF, Mahony D, Lees EM;
XX XX
XX DR WPI; 2000-171015/15.
XX DR N-PSDB; 292404.
XX XX
XX PT New isolated mammalian genes, used to develop products for treating
XX PT e.g. immune, inflammatory or allergic abnormalities, cancers or

```

```

PT degenerative conditions -
XX Claim 24; Page 157; 218pp; English.
XX CC The invention relates to a number of primate and/or rodent proteins, and
XX CC the genes which encode them. The invention encompasses human dendritic
XX CC cell prostaglandin transporter (DC-PGT); the TNF (tumour necrosis
XX CC factor) receptor family-related proteins HDTEA84, HSLJD37R and RANKL;
XX CC human CC chemokine HCC5; human deubiquitinating proteins DUB1 and Dub
XX CC 12; human MD-1 and human and murine MD-2 proteins, which exhibit the
XX CC properties of ligands for proteins comprising a leucine-rich motif (LRR);
XX CC human cyclin E2; cDNAs encoding these proteins; and antibodies against
XX CC these proteins. The proteins can be used for modulating the physiology or
XX CC development of a cell. They can be used to mediate uptake of substrates
XX CC (e.g., prostaglandin-like molecules), to modulate or mediate cellular
XX CC interactions (e.g., induce or prevent trafficking, proliferation, or
XX CC differentiation of cells), or are intracellular proteins which are
XX CC important in various cellular processes such as the deubiquitination of
XX CC proteins or cell cycle regulation. The products can be used for
XX CC treating medical conditions such as immune, inflammatory or allergic
XX CC disorders, or abnormal cellular proliferation, for example, cancers or
XX CC degenerative conditions. They can be used to modulate immune responses in
XX CC disease states e.g., autoimmune disorders, including rheumatoid
XX CC arthritis, systemic lupus erythematosus, Hashimoto's autoimmune
XX CC thyroiditis, as well as acute and chronic inflammatory responses in which
XX CC T cell activation, expansion, and/or immunological T cell memory play an
XX CC important role. Sequences Y77458-Y77461 and Y77465-Y77468 represent
XX CC TNF receptor family-related proteins. Y77458 is the human protein
XX CC HDTEA84, Y77459-Y77461 are human HSLJD37R proteins, Y77465 is
XX CC murine Rank-like protein RANKL, and Y77466-Y77468 are human RANKL
XX CC proteins.
XX SQ Sequence 300 AA:
XX Query Match 100.0%; Score 55; DB 21; Length 300;
XX Best Local Similarity 100.0%; Pred. No. 0.0086;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 GSSSHDTLCT 10
XX DB |||||||||
XX DB 185 gssshdtlct 194
XX
XX RESULT 15
XX ID Y42183 standard; Protein: 302 AA.
XX AC Y42183;
XX XX
XX DT 17-DEC-1999 (first entry)
XX DE Human FLINT #2 protein sequence.
XX XX
XX KW Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
XX KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
XX KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
XX KW reperfusion-associated injury; aplastic anaemia; differentiation;
XX KW growth; myelodysplastic syndrome; pancytopenic condition;
XX KW myocardial ischaemia.
XX XX
XX OS Homo sapiens.
XX XX
XX PN WO9950413-A2.
XX XX
XX PD 07-OCT-1999.
XX XX
XX PF 30-MAR-1999; 99WO-US06797.
XX XX
XX PR 30-MAR-1998; 98US-0079856.
XX PR 20-MAY-1998; 98US-0086074.
XX PR 09-SEP-1998; 98US-0099643.
XX PR 17-DEC-1998; 98US-0112577.

```

```

PR 18-DEC-1998; 98US-0112703.
PR 18-DEC-1998; 98US-0112933.
PR 22-DEC-1998; 98US-0113407.
XX
XX (ELIL ) LILLY & CO ELI.
XX
XX Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
PI Song HY, Wang J, Wu X, Zuckerman SH;
XX
XX WPI: 1999-591319/50.
DR N-PSDB; 225376.
XX
XX
XX Use of mature FLINT for treating acute liver failure, inflammation,
PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
PT and proinflammatory activity
XX
XX
XX Example 2; Fig 2; 99pp; English.
XX
XX The present invention describes therapeutic applications of mature FLINT
CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
CC (mFLINT), which is a member of the tumour necrosis factor receptor
CC superfamily, is used for treating acute liver failure, inflammation of
CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
CC injury or disorder such as hypercoagulation (including use with
CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
CC or disorder, Type I diabetes, cancer, cell damage or damage to an
CC innocent bystander tissue that is induced by a chemotherapeutic agent or
CC therapeutic irradiation, treating haematopoietic progenitor cells that
CC have been exposed to therapeutic radiation or chemotherapy, aplastic
CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
CC also used for promoting the growth or differentiation of a haematopoietic
CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
CC resulting from abnormal myocardial ischaemia. The present sequence
CC represents human FLINT.
XX
XX Sequence 302 AA:
SQ

```

```

Query Match      100.0%; Score 55; DB 20; Length 302;
Best Local Similarity 100.0%; Pred. No. 0.0087;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSSHDTLCT 10
Db 187 gssshdtlct 196

```

Search completed: January 30, 2001, 16:45:23
Job time: 583 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:50:21 ; Search time 149.64 Seconds
(without alignments)
4.538 Million cell updates/sec

Title: US-09-518-931-2_COPY_185_194

Perfect score: 55

Sequence: 1 GSSSHDTLCT 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_66:*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	39	70.9	278	2	B65006	hypothetical prote
2	38	69.1	400	2	T45919	hypothetical prote
3	37	67.3	10797	2	T30192	probable peptide s
4	36	65.5	144	2	C82375	mioc protein VC000
5	36	65.5	146	1	C64085	hypothetical prote
6	36	65.5	277	2	A60771	B-cell activation
7	36	65.5	407	2	T22554	hypothetical prote
8	36	65.5	611	2	T20501	hypothetical prote
9	36	65.5	638	2	T37309	flr-1 protein - Ca
10	36	65.5	1245	2	T18211	delta endotoxin -
11	36	65.5	1437	2	S07430	M polyprotein prec
12	35	63.6	66	2	A29386	neurotoxin III - s
13	35	63.6	419	2	I55463	K+ channel beta-su
14	35	63.6	481	2	T01850	UTP-glucose glucos
15	35	63.6	1205	1	A38943	nitric-oxide synth
16	34	61.8	471	2	S18446	variant surface gl
17	34	61.8	496	2	S37379	catalase (EC 1.11.
18	34	61.8	707	2	S77094	glycogen operon pr
19	34	61.8	1023	2	T31669	neural zinc finger
20	34	61.8	1104	2	I38869	transcription fact
21	33	60.0	225	2	T35686	hypothetical prote
22	33	60.0	241	2	T30761	hypothetical prote
23	33	60.0	243	2	T47410	leucine-rich repea
24	33	60.0	267	2	T46202	endoxylglucan tra
25	33	60.0	289	2	A46515	B cell-associated
26	33	60.0	305	2	A46476	B cell-associated
27	33	60.0	316	2	B70768	hypothetical prote
28	33	60.0	321	2	C64941	hypothetical prote
29	33	60.0	330	2	I39194	gene HOXA1 protein

RESULT 1
B65006
hypothetical protein b2332 - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 18-Aug-2000
C:Accession: B65006
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: B65006
A:Status: Preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-278 <BLAT>
A:Cross-references: GB:AE000322; GB:U00096; NID:gl788672; PIDN:AA75392.1; PID:gl7886
A:Experimental source: strain K-12, substrain MG1655
C:Superfamily: Escherichia coli hypothetical protein b2332

Query Match 70.9%; Score 39; DB 2; Length 278;
Best Local Similarity 66.7%; Pred. No. 3.1;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GSSSHDTLCT 9
| | | | |
DB 150 GPDSDHTC 158

RESULT 2
T45919
hypothetical protein F5K20.90 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C:Accession: T45919
R:Monfort, A.; Casacuberta, E.; Puigdomenech, P.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23017
A:Accession: T45919
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-400 <MON>
A:Cross-references: EMBL:AL132960
A:Experimental source: cultivar Columbia; BAC clone F5K20
C:Genetics:
A:Map position: 3
A:Introns: 22/3; 127/1; 173/1; 277/2; 337/2; 358/3
A>Note: F5K20.90

Query Match 69.1%; Score 38; DB 2; Length 400;
Best Local Similarity 77.8%; Pred. No. 7.2;

Wed Jan 31 13:44:00 2001

```
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 SSSHDITLCT 10
   | | | | | | |
Db 164 SESHDTPCT 172

RESULT 3
T30192
C:Species: Aureobasidium pullulans
C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
C:Accession: T30192
R:Peery, R.B.; Thornwell, S.J.; Tobin, M.B.; Skatrud, P.L.
submitted to the EMBL Data Library, January 1997
A:Description: Discovery of an MDR-like gene adjacent to a peptide synthetase in Aureoba
A:Reference number: Z20767
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-10797 <PEE>
A:Cross-references: EMBL:U85909; NID:g4099310; PID:g4099313; PIDN:AAD00581.1
C:Genetics:
A:Introns: 2078/3; 4142/3; 6075/3; 7963/3; 7985/2

Query Match 67.3%; Score 37; DB 2; Length 10797;
Best Local Similarity 75.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 SSSHDITLCT 10
   : | | | | | | |
Db 9444 ASHDVLCT 9451

RESULT 4
mioc protein VC0002 [imported] - Vibrio cholerae (group O1 strain N16961)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Sep-2000
C:Accession: C82375
R:Heidelber, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B.
I, R.K.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: AB2035; MUID:20406833
A:Accession: C82375
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-144 <HEI>
A:Cross-references: GB:AE004093; GB:AE003852; NID:g9654391; PIDN:AAF93180.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC0002
A:Map position: 1
C:Superfamily: mioc protein; flavodoxin homology
C:Keywords: flavoprotein

Query Match 65.5%; Score 36; DB 2; Length 144;
Best Local Similarity 66.7%; Pred. No. 6.3;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GSSSHDTLCT 9
   | | | | | | |
Db 89 GDSSYDTFC 97

RESULT 5
C64085
C:Species: Haemophilus influenzae (strain Rd KW20)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 28-Jul-2000
C:Accession: C64085
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman
, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vente
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630
A:Accession: C64085
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-146 <TTGR>
A:Cross-references: GB:U32750; GB:I42023; NID:gl573668; PIDN:AAC22329.1; PID:gl573670
C:Superfamily: mioc protein; flavodoxin homology
C:Keywords: flavoprotein
F:5-141/Domain: flavodoxin homology <FLX>

Query Match 65.5%; Score 36; DB 1; Length 146;
Best Local Similarity 66.7%; Pred. No. 6.4;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GSSSHDTLCT 9
   | | | | | | |
Db 90 GSSYDTFC 98

RESULT 6
A60771
B-cell activation protein CD40 precursor - human
N:Alternate names: B-cell surface antigen Bp50
C:Species: Homo sapiens (man)
C:Date: 03-Jun-1993 #sequence_revision 03-Feb-1994 #text_change 21-Jul-2000
C:Accession: S04460; A60771
R:Stamenkovic, I.; Clark, E.A.; Seed, B.
EMBO J. 8, 1403-1410, 1989
A:Title: A B-lymphocyte activation molecule related to the nerve growth factor recept
A:Reference number: S04460; MUID:89356608
A:Accession: S04460
A:Molecule type: mRNA
A:Residues: 1-277 <STA>
A:Cross-references: EMBL:X60592; NID:g29850; PIDN:CAA43045.1; PID:g29851
R:Brasch-Andersen, S.; Paulie, S.; Koho, H.; Nika, H.; Aspenstroem, P.; Perlmann, P.
J. Immunol. 142, 562-567, 1989
A:Title: Biochemical characteristics and partial amino acid sequence of the receptor-
A:Reference number: A60771; MUID:89093941
A:Accession: A60771
A:Molecule type: protein
A:Residues: 21-50 <BRA>
A:Experimental source: Burkitt lymphoma cell line Raji
C:Genetics:
A:Gene: GDB:CD40
A:Cross-references: GDB:215268; OMIM:109535
A:Map position: 20q12-20q13.2
C:Superfamily: CD27 antigen; NGF receptor repeat homology
C:Keywords: B-cell; glycoprotein; phosphoprotein; surface antigen; transmembrane prot
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-277/Product: B-cell activation protein CD40 #status experimental <MAT>
F:21-193/Domain: extracellular #status predicted <EXT>
F:194-215/Domain: transmembrane #status predicted <TM>
F:216-277/Domain: intracellular #status predicted <CYT>
F:153,180/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 65.5%; Score 36; DB 2; Length 277;
Best Local Similarity 60.0%; Pred. No. 12;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GSSSHDTLCT 10
   | | | | | | |
Db 95 GTSYDTLCT 104
```

```
RESULT 7
T22554
hypothetical protein F53C11.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T22554
R:Baynes, C.
submitted to the EMBL Data Library, September 1996
A:Reference number: Z19581
A:Accession: T22554
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-407 <WIL>
A:Cross-references: EMBL:Z79756; PIDN:CA02116.1; GSPDB:GN00023; CESP:F53C11.7
A:Experimental source: clone F53C11
C:Genetics:
A:Gene: CESP:F53C11.7
A:Map position: 5
A:Introns: 61/3; 134/1; 162/3; 215/3; 311/3; 380/3

Query Match 65.5%; Score 36; DB 2; Length 407;
Best Local Similarity 70.0%; Pred. No. 18;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GSSSHDTLCT 10
I:|:|:|:|:|
Db 201 GTSSIDTCT 210

RESULT 8
T20501
hypothetical protein F02D10.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T20501
R:Swinburne, J.
submitted to the EMBL Data Library, November 1995
A:Reference number: Z19283
A:Accession: T20501
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-611 <WIL>
A:Cross-references: EMBL:Z67990; PIDN:CAA91936.1; GSPDB:GN00028; CESP:F02D10.5
A:Experimental source: clone F02D10
C:Genetics:
A:Gene: CESP:F02D10.5
A:Map position: X
A:Introns: 21/3; 61/3; 86/3; 111/1; 150/3; 190/2; 222/1; 268/3; 292/3; 344/1; 355/1; 369/1

Query Match 65.5%; Score 36; DB 2; Length 611;
Best Local Similarity 66.7%; Pred. No. 28;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SSSHDTLCT 10
I:|:|:|:|:|
Db 323 SGSNDTICT 331

RESULT 9
T37309
flr-1 protein - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
C:Accession: T37309
R:Takeuchi, M.; Kawakami, M.; Ishihara, T.; Amano, T.; Kondo, K.; Katsura, I.
Proc. Natl. Acad. Sci. U.S.A. 95, 11775-11780, 1998
A:Title: An ion channel of the degenerin/epithelial sodium channel superfamily controls
A:Reference number: Z21675; MUID:98426229
A:Accession: T37309
A:Status: preliminary; translated from GB/EMBL/DBJ
```

```
A:Molecule type: mRNA
A:Residues: 1-638 <TAK>
A:Cross-references: EMBL:AB012617; NID:g3702221; PIDN:BA033488.1; PID:g3702222
C:Genetics:
A:Note: flr-1

Query Match 65.5%; Score 36; DB 2; Length 638;
Best Local Similarity 66.7%; Pred. No. 29;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SSSHDTLCT 10
I:|:|:|:|:|
Db 333 SGSNDTICT 341

RESULT 10
T18211
delta endotoxin - Bacillus thuringiensis (fragment)
C:Species: Bacillus thuringiensis
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18211
R:Foncerrada, L.; Narva, K.E.
submitted to the EMBL Data Library, January 1995
A:Description: Bacillus thuringiensis PS86Q3 delta endotoxin.
A:Reference number: Z18828
A:Accession: T18211
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1245 <FON>
A:Cross-references: EMBL:U19725; NID:g862636; PID:g862637; PIDN:AAA68598.1
C:Keywords: delta-endotoxin

Query Match 65.5%; Score 36; DB 2; Length 1245;
Best Local Similarity 77.8%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SSSHDTLCT 10
I:|:|:|:|:|
Db 773 SSEHDTLAT 781

RESULT 11
S07430
M polyprotein precursor - Germiston virus
N:Contains: glycoprotein G1; glycoprotein G2; nonstructural protein NS-M
C:Species: Germiston virus
C>Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Sep-1994
C:Accession: S07430
R:Parignon, N.; Vialat, P.; Gerbaud, S.; Girard, M.; Bouloy, M.
Virus Res. 11, 73-85, 1988
A:Title: Nucleotide sequence of the M segment of Germiston virus: comparison of the M
A:Reference number: S07430; MUID:89021381
A:Accession: S07430
A:Molecule type: genomic RNA
A:Residues: 1-1437 <PAR>
A:Cross-references: EMBL:M21951
C:Genetics:
A:Gene: M
C:Superfamily: bunyavirus M polyprotein
C:Keywords: glycoprotein; nonstructural protein; polyprotein; transmembrane protein
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-1437/Product: M polyprotein #status predicted <MP>
F:1382-1412/Domain: transmembrane #status predicted <TM>
F:65,88,252,474,627,897,1173/Binding site: carbohydrate (Asn) (covalent) #status pred

Query Match 65.5%; Score 36; DB 2; Length 1437;
Best Local Similarity 66.7%; Pred. No. 68;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SSSHDTLCT 10
```

```

C:Genetics:
A:Map position: 4
A:Note: F9D12.4
C:Superfamily: flavonol O3-glucosyltransferase

Query Match      63.6%   Score 35; DB 2; Length 481;
Best Local Similarity 55.6%   Pred. No. 35;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GSSSHDTLC 9
Db 453 GCSAHSILC 461

RESULT 15
A38943
nitric-oxide synthase (EC 1.14.13.39), endothelial - bovine
N:Alternate names: ECNOS; nitric-oxide synthase type III
C:Species: Bos primigenius taurus; (cattle)
C:Date: 02-Jun-1995 #sequence_revision 02-Jun-1995 #text_change 03-Mar-2000
C:Accession: A38943; A46033; I45945; A42841; I45946; A38944
R:Lamas, S.; Marsden, P.A.; Li, G.K.; Tempst, P.; Michel, T.
submitted to GenBank, July 1992
A:Reference number: A38943
A:Accession: A38943
A:Molecule type: mRNA
A:Residues: 1-1205 <LAM1>
A:Cross-references: GB:M89952; NID:g162976; PIDN:AAA30494.1; PID:g162977
A:Experimental source: aortic endothelial cells
R:Lamas, S.; Marsden, P.A.; Li, G.K.; Tempst, P.; Michel, T.
Proc. Natl. Acad. Sci. U.S.A. 89, 6348-6352, 1992
A:Title: Endothelial nitric oxide synthase: molecular cloning and characterization of
A:Reference number: A46033; MUID:92335295
A:Accession: A46033
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-860, 'I', 862-1205 <LAM2>
A:Cross-references: GB:M89952; NID:g162976
A:Experimental source: endothelial
A:Note: sequence extracted from NCBI backbone (NCBIP:108720)
R:Nishida, K.; Harrison, D.G.; Navas, J.P.; Fisher, A.A.; Dockery, S.P.; Nerem, R.M.;
J. Clin. Invest. 90, 2092-2096, 1992
A:Title: Molecular Cloning and Characterization of the constitutive bovine aortic End
A:Reference number: I45945; MUID:93055452
A:Accession: I45945
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1205 <NIS>
A:Cross-references: GB:M99057; NID:g163421; PIDN:AAA30667.1; PID:g163422
R:Sessa, W.C.; Harrison, J.K.; Barber, C.M.; Zeng, D.; Durieux, M.E.; D'Angelo, D.D.;
J. Biol. Chem. 267, 15274-15276, 1992
A:Title: Molecular cloning and expression of a cDNA encoding endothelial cell nitric
A:Reference number: A42841; MUID:92348367
A:Accession: A42841
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-51, 'N', 53-99, 'R', 101-147, 'M', 149-164, 'I', 166-317, 'GA', 320, 'HTGVRGP', 32
'K', 517-692, 'G', 694-740, 'A', 742-753, 'N', 755-799, 'N', 801-803, 'SA', 806-856, 'V', 858-906,
A:Experimental source: aortic endothelial cells
A:Note: sequence extracted from NCBI backbone (NCBIP:109564); contains a number of ty
A:Accession: I45946
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-99, 'R', 101-164, 'I', 166-317, 'GA', 320, 'HTGVRGP', 329-454, 'Y', 456-458, 'P',
A:Cross-references: GB:M95674; NID:g163426; PIDN:AAA30669.1; PID:g163427
A:Experimental source: aortic endothelial cells
A:Note: submitted to GenBank, August 1992
A:Note: GenBank entry BOVNOS, release 103.0, has a typographical error in the referen
C:Function:
A:Description: catalyzes the oxidation of an L-arginine guanidino nitrogen and of NAD
C:Superfamily: nitric-oxide synthase; flavodoxin homology; NADPH-ferrihemoprotein re
C:Keywords: blocked amino end; calmodulin binding; chromoprotein; FAD; flavoprotein;

Db 1030 NSKHDELCT 1038

RESULT 12
A29386
neurotoxin III - scorpion (Buthus occitanus)
C:Species: Buthus occitanus mardochei
C:Date: 21-May-1988 #sequence_revision 21-May-1988 #text_change 07-Feb-1997
C:Accession: A29386
R:Vargas, O.; Martin, M.F.; Rochat, H.
Eur. J. Biochem. 162, 589-599, 1987
A:Title: Characterization of six toxins from the venom of the Moroccan scorpion Buthus o
A:Reference number: A29386; MUID:87161829
A:Accession: A29386
A:Molecule type: protein
A:Residues: 1-66 <VAR>
C:Superfamily: scorpion neurotoxin
C:Keywords: neurotoxin

Query Match      63.6%   Score 35; DB 2; Length 66;
Best Local Similarity 77.8%   Pred. No. 4.5;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GSSSHDTLC 9
Db 19 GSSGCDTLC 27

RESULT 13
I55463
K+ channel beta-subunit - human
C:Species: Homo sapiens (man)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 18-Feb-2000
C:Accession: I55463
R:England, S.K.; Uebale, V.N.; Kodali, J.; Bennett, P.B.; Tamkun, M.M.
J. Biol. Chem. 270, 28531-28534, 1995
A:Title: A novel K+ channel beta-subunit (hkv beta 1.3) is produced via alternative mRNA
A:Reference number: I55463; MUID:96081908
A:Accession: I55463
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-419 <RES>
A:Cross-references: GB:I47665; NID:g1088278; PIDN:AAC41926.1; PID:g1088279
C:Superfamily: fission yeast pyridoxine 4-dehydrogenase

Query Match      63.6%   Score 35; DB 2; Length 419;
Best Local Similarity 66.7%   Pred. No. 30;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SSSHTLTCT 10
Db 78 SSEHTTCT 86

RESULT 14
T01850
UTP-glucose glucosyltransferase homolog F9D12.4 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 24-Nov-1999
C:Accession: T01850
R:Murray, J.; Langston, Y.; Ahrens, C.
submitted to the EMBL Data Library, July 1998
A:Description: The sequence of Arabidopsis thaliana F9D12.
A:Reference number: Z14444
A:Accession: T01850
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-481 <MUR>
A:Cross-references: EMBL:AF077407; NID:g3319339; PID:g3319344
A:Experimental source: cultivar Columbia

```

F:493-512/Region: calmodulin binding #status predicted
F:522-1161/Domain: NADPH--ferrihemoprotein reductase homology <FEH>
F:524-705/Domain: flavodoxin homology <FLX>
F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
F:3/Modified site: aspartic acid (Asn) #status predicted
F:186/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 63.6%; Score 35; DB 1; Length 1205;
Best Local Similarity 66.7%; Pred. No. 89;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GSSSHDTLC 9
|| :|||
Db 79 GSITYDTLC 87

Search completed: January 30, 2001, 16:50:24
Job time: 714 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 17:05:57 ; Search time 80.04 Seconds
(without alignments)
4.035 Million cell updates/sec

Title: US-09-518-931-2_COPY_185_194
Perfect score: 55
Sequence: 1 GSSSHDTLCT 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	39	70.9	273	1 YFCO_ECOLI	P76498 escherichia
2	36	65.5	146	1 MIOC_HAEIN	P44813 haemophilus
3	36	65.5	277	1 CD40_HUMAN	P25942 homo sapien
4	36	65.5	1245	1 C5BA_BACTU	Q45712 bacillus th
5	36	65.5	1437	1 VGLM_BUNGE	P12430 bunyavirus
6	35	63.6	66	1 SCX3_BUTOM	P13488 buthus occi
7	35	63.6	67	1 SCL3_LEIQH	P56678 leirus qui
8	35	63.6	1204	1 NOS3_BOVIN	P29473 bos taurus
9	35	63.6	1204	1 NOS3_PIG	Q28969 sus scrofa
10	34	61.8	471	1 VS11_TRYBB	P26326 trypanosoma
11	34	61.8	496	1 CAT3_MAI2E	P18123 zea mays (m
12	34	61.8	1104	1 NFXL_HUMAN	Q12986 homo sapien
13	33	60.0	241	1 CELA_MCV1	Q98325 molluscum c
14	33	60.0	289	1 CD40_MOUSE	P27512 mus musculu
15	33	60.0	316	1 YK95_MYCTU	Q10704 mycobacteri
16	33	60.0	321	1 YEAX_ECOLI	P76254 escherichia
17	33	60.0	335	1 HXAL_HUMAN	P49639 homo sapien
18	33	60.0	380	1 MEDH_BACWT	P31005 bacillus me
19	33	60.0	434	1 FES_ERWCH	O51900 erwinia chr
20	33	60.0	461	1 TNR2_HUMAN	P20333 homo sapien
21	33	60.0	507	1 LCK_CHICK	P42683 gallus gall
22	33	60.0	719	1 PRL1_SCHPO	Q03319 schizosacch
23	33	60.0	1239	1 POLS_EEEV	P08768 eastern equ
24	33	60.0	1240	1 POLS_EEEV3	P27284 eastern equ
25	33	60.0	1441	1 VGLM_BUNL7	P09612 bunyavirus
26	33	60.0	1441	1 VGLM_BUNSH	P04875 bunyavirus
27	33	60.0	1455	1 FACA_HUMAN	O15360 homo sapien
28	33	60.0	1551	1 VGLM_DUGBV	Q02004 dugbe virus
29	32.5	59.1	864	1 SBE2_YEAST	P42223 saccharomyc
30	32	58.2	127	1 WAP_RABIT	P09412 cryptotolagu
31	32	58.2	247	1 YK95_MYCLE	P54075 mycobacteri
32	32	58.2	280	1 YBB0_YEAST	P32788 saccharomyc
33	32	58.2	374	1 FOS_TETFL	Q91496 tetraodon f

34	32	58.2	390	1 DAF1_MOUSE	O61475 mus musculu
35	32	58.2	395	1 INX3_DROME	O9vas7 drosophila
36	32	58.2	407	1 DAF2_MOUSE	O61476 mus musculu
37	32	58.2	512	1 Y4WA_RHISN	P55679 rhizobium s
38	32	58.2	550	1 GYRA_MYCGO	Q49467 mycobacteri
39	32	58.2	560	1 INR1_BOVIN	Q04790 bos taurus
40	32	58.2	560	1 INR1_SHEEP	Q28589 ovis aries
41	32	58.2	707	1 YJ9C_YEAST	P47166 saccharomyc
42	32	58.2	914	1 NRPL_CHICK	P79795 gallus gall
43	31	56.4	67	1 SCX3_ORTSC	P15225 orthochirur
44	31	56.4	146	1 MIOC_ECOLI	P03817 escherichia
45	31	56.4	248	1 YJHP_ECOLI	P39367 escherichia

ALIGNMENTS

RESULT 1
YFCO_ECOLI
AC P76498: STANDARD; PRT; 273 AA.
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HYPOTHETICAL 30.3 KDA PROTEIN IN AROC-SIXA INTERGENIC REGION
DE PRECURSOR.
GN YFCO.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE-97426617; PubMed-9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
CC -!- SIMILARITY: STRONG, TO S.TYPHIMURIUM YADU.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF000322; AAC75392.1; ALT_INIT.
DR ECOGENE: EGI4118; YFCO.
KW Hypothetical protein: Signal.
FT SIGNAL 1 21
FT CHAIN 22 273 HYPOTHETICAL PROTEIN YFCO.
SQ SEQUENCE 273 AA; 30260 MW; 9D1D9C22959823DC CRC64;

Query Match 70.9%; Score 39; DB 1; Length 273;
Best Local Similarity 66.7%; Pred. No. 1.6;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GSSSHDTLCT 9
| | | | |
DB 145 GPDSDHTIC 153

RESULT 2
MIOC_HAEIN STANDARD; PRT; 146 AA.
ID MIOC_HAEIN
AC P44813;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)

01-NOV-1995 (Rel. 32, Last annotation update)
DE MIOC PROTEIN HOMOLOG.
GN MIOC OR HI0669
OS Haemophilus influenzae
OC Bacteria: Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
[1]
RN SEQUENCE FROM N.A.
RC STRAIN-RD / KW20:
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.";
RL Science 269:496-512(1995).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U32750; AAC22329.1; -
DR HSSP; P00323; IFX1.
DR TIGR; HI0669; -
SQ SEQUENCE 146 AA; 16030 MW; 2A08A4E4133953EA CRC64;

Query Match 65.5%; Score 36; DB 1; Length 146;
Best Local Similarity 66.7%; Pred. No. 3.2;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GSSSHDTLC 9
 | | | | |
DB 90 GSSDYDTFC 98

RESULT 3
CD40_HUMAN STANDARD; PRT; 277 AA.
ID CD40_HUMAN
AC P25942;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE CD40L RECEPTOR PRECURSOR (B-CELL SURFACE ANTIGEN CD40) (BP50) (CDW40)
DE (TUMOR NECROSIS FACTOR RECEPTOR 5).
GN TNFRSF5 OR CD40.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
[1]
RN SEQUENCE FROM N.A.
RC MEDLINE=89356608; PubMed=2475341;
RX Stamenkovic I., Clark E.A., Seed B.;
RA "A B-lymphocyte activation molecule related to the nerve growth
RT factor receptor and induced by cytokines in carcinomas.";
RL EMBO J. 8:1403-1410(1989).
[2]
RN 3D-STRUCTURE MODELING OF 24-144.
RX MEDLINE=97189482; PubMed=9037712;
RA Bajorath J., Aruffo A.;
RT "Construction and analysis of a detailed three-dimensional model of
RT the ligand binding domain of the human B cell receptor CD40.";
RL Proteins 27:59-70(1997).

[3]
3D-STRUCTURE MODELING OF 26-186 IN COMPLEX WITH CD40L.
RX MEDLINE=98266353; PubMed=9605317;
RA Singh J., Garber E., van Vlijmen H., Karpusas M., Hsu Y.-M.,
RA Zheng Z., Naismith J.H., Thomas D.;
RT "The role of polar interactions in the molecular recognition of CD40L
RT with its receptor CD40.";
RL Protein Sci. 7:1124-1135(1998).
CC -1- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD40L.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: B-CELLS AND IN PRIMARY CARCINOMAS.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD40 entry; "WWW=<http://www.ncbi.nlm.nih.gov/prow/cd/cd40.htm>".

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X60592; CAA43045.1; -
DR PIR; S04460; S04460.
DR PDB; LCDF; 01-APR-97.
DR MIM; 109535; -
DR INTERPRO; IPR001368; -
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 1.
DR PROSITE; PS50050; TNFR_NGFR_2; 4.
KW Receptor; B-cell; Glycoprotein; Transmembrane; Repeat; Signal;
3D-structure.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 277 CD40L RECEPTOR.
FT DOMAIN 20 193 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 194 215 POTENTIAL.
FT DOMAIN 216 277 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 25 187 4 X TNFR-CYS.
FT REPEAT 25 60 TNFR-CYS 1.
FT REPEAT 61 103 TNFR-CYS 2.
FT REPEAT 104 144 TNFR-CYS 3.
FT REPEAT 145 187 TNFR-CYS 4.
FT DISULFID 26 37
FT DISULFID 38 51
FT DISULFID 41 59
FT DISULFID 62 77
FT DISULFID 83 103
FT DISULFID 105 119
FT DISULFID 111 116
FT DISULFID 125 143
FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 277 AA; 30619 MW; BC8776EC2CA5680 CRC64;

Query Match 65.5%; Score 36; DB 1; Length 277;
Best Local Similarity 60.0%; Pred. No. 6.2;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GSSSHDTLC 10
 | | | | |
DB 95 GTSETDTICT 104

RESULT 4
CSBA_BACTU STANDARD; PRT; 1245 AA.
ID CSBA_BACTU
AC Q45712;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PESTICIDIAL CRYSTAL PROTEIN CRY5BA (INSECTICIDAL DELTA-ENDOTOXIN
DE

DE CRVB(A)) (CRYSTALLINE ENTOMOCIDAL PROTOXIN) (140 KDA CRYSTAL PROTEIN).
 GN CRYSBA OR CRVB(A).
 OS Bacillus thuringiensis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 CC Bacillus/Staphylococcus group; Bacillus.
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-15.
 RC STRAIN-NRRL B-18765 / PS86Q3;
 RA Payne J.M., Kennedy M.K., Randall J.B., Meier H., Uick H.J.,
 RA Foncarrada L., Schnepf H.E., Schwab G.E., Fu J.M.;
 RT "Bacillus thuringiensis toxins active against hymenopteran pests.";
 RL Patent number US5596071, 21-JAN-1997.
 CC -1- FUNCTION: PROMOTES COLLOIDOSMOTIC LYSIS BY BINDING TO THE MIDGUT
 CC EPITHELIAL CELLS OF HYMENOPTERAN SPECIES.
 CC -1- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING
 CC SPOULATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART
 CC OF THE SPORE COAT.
 CC -1- MISCELLANEOUS: TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE
 CC N-TERMINUS.
 CC -1- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U19725; AAA68598.1; -;
 DR INTERPRO; IPR001178; -;
 DR PFAM; PF00555; endotoxin; 1.
 KW Toxin; Sporulation.
 SQ SEQUENCE 1245 AA; 139775 MW; C7CB93F96650171F CRC64;

Query Match 65.5%; Score 36; DB 1; Length 1245;
 Best Local Similarity 77.8%; Pred. No. 30;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SSSHDTLCT 10
 || |||||
 Db 773 SSEHDTLAT 781

RESULT 5
 VGLM_BUNGE STANDARD; PRT; 1437 AA.
 ID VGLM_BUNGE STANDARD; PRT; 1437 AA.
 AC P12430;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-AUG-1992 (Rel. 23, Last annotation update)
 DE M POLYPROTEIN PRECURSOR [CONTAINS: NONSTRUCTURAL PROTEIN NS-M;
 DE M POLYPROTEINS G1 AND G2].
 OS Bunyavirus germiston.
 GN Bunyavirus.
 OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Bunyavirus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-89021381; PubMed=3176688;
 RA Pardigon N., Vialat P., Gerbaud S., Girard M., Bouloy M.;
 RT "Nucleotide sequence of the M segment of Germiston virus: comparison
 RT of the M gene product of several bunyaviruses.";
 RL Virus Res. 11:73-85(1988).
 CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS
 CC INCLUDING NONSTRUCTURAL PROTEIN NS-M, GLYCOPROTEIN G1, AND
 CC GLYCOPROTEIN G2.
 CC -1- SIMILARITY: BELONGS TO THE BUNYAVIRUSES M POLYPROTEIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M21951; AAA42778.1; -;
 DR PIR; S07430; S07430.
 KW Polyprotein; Glycoprotein; Transmembrane; Nonstructural protein;
 CC Signal.
 FT SIGNAL 1 21 M POLYPROTEIN.
 FT CHAIN 22 1437 GLYCOPROTEIN G2.
 FT CHAIN 22 306 NONSTRUCTURAL PROTEIN NS-M.
 FT CHAIN 307 481 GLYCOPROTEIN G1.
 FT CHAIN 482 1437 POTENTIAL.
 FT TRANSMEM 208 228 POTENTIAL.
 FT TRANSMEM 316 336 POTENTIAL.
 FT TRANSMEM 368 388 POTENTIAL.
 FT TRANSMEM 1395 1415 POTENTIAL.
 FT CARBOHYD 65 65 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 88 88 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 252 252 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 474 474 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 627 627 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1173 1173 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 1437 AA; 162493 MW; C4FB3001BD09A30D CRC64;

Query Match 65.5%; Score 36; DB 1; Length 1437;
 Best Local Similarity 66.7%; Pred. No. 35;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SSSHDTLCT 10
 :| |||||
 Db 1030 NSKHDELCT 1038

RESULT 6
 SCX3_BUTOM STANDARD; PRT; 66 AA.
 ID SCX3_BUTOM STANDARD; PRT; 66 AA.
 AC P13488;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE NEUROTOXIN III (BOM III).
 OS Buthus occitanus mardochel (Moroccan scorpion).
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
 OC Buthoidea; Buthidae; Buthus.
 RN [1]
 RP SEQUENCE.
 RX MEDLINE-87161829; PubMed=3104036;
 RA Vargas O., Martin M.-F., Rochat H.;
 RT "Characterization of six toxins from the venom of the Moroccan
 RT scorpion Buthus occitanus mardochel.";
 RL Eur. J. Biochem. 162:589-599(1987).
 CC -1- FUNCTION: BINDS TO SODIUM CHANNELS AND INHIBIT THE INACTIVATION OF
 CC THE ACTIVATED CHANNELS, THEREBY BLOCKING NEURONAL TRANSMISSION.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
 CC ALPHA-TOXIN SUBFAMILY.
 DR PIR; A29386; A29386.
 DR HSSP; P01487; ILQQ.
 DR INTERPRO; IPR002061; -;
 DR PFAM; PF00537; Toxin_3; 1.
 KW Neurotoxin; Sodium channel inhibitor.
 FT DISULFID 12 65 BY SIMILARITY.
 FT DISULFID 16 37 BY SIMILARITY.
 FT DISULFID 23 47 BY SIMILARITY.
 FT DISULFID 27 49 BY SIMILARITY.
 SQ SEQUENCE 66 AA; 6872 MW; 7A96DA393F89852F CRC64;

Query Match 63.6%; Score 35; DB 1; Length 66;
 Best Local Similarity 77.8%; Pred. No. 2.1;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT

or send an email to license@isb-sib.ch.

```

CC EMBL; M99057; AAA30667.1; -;
CC EMBL; M89952; AAA30494.1; -;
CC EMBL; M95674; AAA30669.1; -;
CC PDB; 1NSE; 18-MAY-99.
CC PDB; 2NSE; 25-MAY-99.
CC PDB; 3NSE; 18-MAY-99.
CC PDB; 4NSE; 18-MAY-99.
CC INTERPRO; IPR001094; -;
CC INTERPRO; IPR001433; -;
CC INTERPRO; IPR001709; -;
CC INTERPRO; IPR003097; -;
CC PFAM; PF00667; FAD_binding; 1.
CC PRINTS; PR00369; FLAVODOXIN.
CC PRINTS; PR00371; FPNCR.
KW Oxidoreductase; NADP; FAD; FMN; Calmodulin-binding; Myristate;
KW Lipoprotein; Palmitate; Phosphorylation; Calcium-binding; Heme;
KW Zinc; Metal-binding; Multigene family; 3D-structure.
FT INIT_MET 0
FT BINDING 185 185 HEME (BY SIMILARITY).
FT DOMAIN 491 511 CALMODULIN-BINDING (POTENTIAL).
FT NP_BIND 650 681 FMN (PYRIMIDINE PART) (BY SIMILARITY).
FT NP_BIND 794 805 FAD (ADP PART) (BY SIMILARITY).
FT NP_BIND 936 946 FAD (FLAVIN PART) (BY SIMILARITY).
FT NP_BIND 1011 1029 NADP (RIBOSE PART) (BY SIMILARITY).
FT NP_BIND 1109 1124 NADP (ADP PART) (BY SIMILARITY).
FT LIPID 1 1 MYRISTATE.
FT LIPID 14 14 PALMITATE.
FT LIPID 25 25 PALMITATE.
FT METAL 95 95 ZINC.
FT METAL 100 100 ZINC.
FT MOD_RES 142 142 PHOSPHORYLATION (BY CAPK).
FT CONFLICT 99 99 C -> R (IN REF. 3).
FT CONFLICT 164 164 Y -> I (IN REF. 3).
FT CONFLICT 317 327 EHPTLEWFAAL -> GAPHTGWRGP (IN REF. 3).
FT CONFLICT 454 454 S -> Y (IN REF. 3).
FT CONFLICT 458 458 T -> P (IN REF. 3).
FT CONFLICT 740 740 T -> A (IN REF. 3).
FT CONFLICT 803 804 CP -> SA (IN REF. 3).
FT CONFLICT 856 856 L -> V (IN REF. 3).
FT CONFLICT 906 907 WF -> LV (IN REF. 3).
FT CONFLICT 1041 1041 A -> H (IN REF. 3).
SQ SEQUENCE 1204 AA; 133155 MW; D017210062ABE4B0 CRC64;

Query Match 63.6%; Score 35; DB 1; Length 1204;
Best Local Similarity 66.7%; Pred. NO. 46;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSSSDHTLC 9
DB 78 GSITYDTLC 86

RESULT 9
ID NOS3_PIG STANDARD; PRT; 1204 AA.
AC Q28969;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE NITRIC-OXIDE SYNTHASE, ENDOTHELIAL (EC 1.14.13.39) (EC-NOS) (NOS, TYPE
DE III) (NOSIII) (ENDOTHELIAL NOS) (ENOS) (CONSTITUTIVE NOS) (CNOS).
GN NOS3 OR NOS.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
[1]
RN SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC TISSUE=PULMONARY ARTERY;
RX MEDLINE=97293429; PubMed=9149402;

```

```

RA Zhang J., Patel J.M., Block E.R.;
RT "Molecular cloning, characterization and expression of a nitric oxide
RL synthase from porcine pulmonary artery endothelial cells.";
RN Comp. Biochem. Physiol. 116B:485-491(1997).
RP [2]
SEQUENCE OF 1031-1205 FROM N.A.
RA Patel J.M., Block E.R.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PRODUCES NITRIC OXIDE (NO) WHICH IS IMPLICATED IN
CC VASCULAR SMOOTH MUSCLE RELAXATION THROUGH A GMP-MEDIATED SIGNAL
CC TRANSDUCTION PATHWAY. NO MEDIATES VASCULAR ENDOTHELIAL GROWTH
CC FACTOR (VEGF)-INDUCED ANGIOGENESIS IN CORONARY VESSELS AND
CC PROMOTES BLOOD CLOTTING THROUGH THE ACTIVATION OF PLATELETS (BY
CC SIMILARITY).
CC -!- CATALYTIC ACTIVITY: L-ARGININE + N NADPH + M O(2) - CITRULLINE +
CC NITRIC OXIDE + N NADP(+).
CC -!- COFACTOR: HEME. BINDS ONE MOLE EACH OF FAD AND FMN. ALSO REQUIRES
CC TETRAHYDROBIOPTERIN (BH4) WHICH MAY STABILIZE THE DIMERIC FORM OF
CC THE ENZYME (BY SIMILARITY).
CC -!- ENZYME REGULATION: STIMULATED BY CALCIUM/CALMODULIN (BY
CC SIMILARITY).
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- INDUCTION: REPRESSED BY PROINFLAMMATORY CYTOKINES.
CC -!- SIMILARITY: BELONGS TO THE NOS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; U59924; AAB39539.1; -;
DR EMBL; U33832; AAB84933.1; -;
DR HSSP; P29477; 2NOS.
DR INTERPRO; IPR001094; -;
DR INTERPRO; IPR001433; -;
DR INTERPRO; IPR001709; -;
DR INTERPRO; IPR003097; -;
DR PFAM; PF00667; FAD_binding; 1.
DR PFAM; PF00175; Oxidored_fad; 1.
DR PRINTS; PR00369; FLAVODOXIN.
DR PRINTS; PR00371; FPNCR.
KW Oxidoreductase; NADP; FAD; FMN; Calmodulin-binding; Myristate;
KW Lipoprotein; Palmitate; Calcium-binding; Heme; Zinc; Metal-binding;
KW Multigene family.
FT INIT_MET 0
FT BINDING 185 185 HEME (BY SIMILARITY).
FT DOMAIN 491 511 CALMODULIN-BINDING (POTENTIAL).
FT NP_BIND 650 681 FMN (PYRIMIDINE PART) (BY SIMILARITY).
FT NP_BIND 794 805 FAD (ADP PART) (BY SIMILARITY).
FT NP_BIND 936 946 FAD (FLAVIN PART) (BY SIMILARITY).
FT NP_BIND 1011 1029 NADP (RIBOSE PART) (BY SIMILARITY).
FT NP_BIND 1109 1124 NADP (ADP PART) (BY SIMILARITY).
FT LIPID 1 1 MYRISTATE (BY SIMILARITY).
FT LIPID 14 14 PALMITATE (BY SIMILARITY).
FT LIPID 25 25 PALMITATE (BY SIMILARITY).
FT METAL 95 95 ZINC (BY SIMILARITY).
FT METAL 100 100 ZINC (BY SIMILARITY).
SQ SEQUENCE 1204 AA; 133274 MW; 48676BA95D814CFE CRC64;

Query Match 63.6%; Score 35; DB 1; Length 1204;
Best Local Similarity 66.7%; Pred. NO. 46;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSSSDHTLC 9
DB 78 GSITYDTLC 86

RESULT 10

```

RA Abler M.L., Scandalios J.G.:
RT "Isolation and characterization of a genomic sequence encoding the
RL maize Cat3 catalase gene."
RN Plant Mol. Biol. 22:1031-1038(1993).
RP [2]
RC SEQUENCE FROM N.A.
RX STRAIN=CV. W64A; TISSUE=EPICOTYL;
RA MEDLINE=89051000; PubMed=2461221;
RT Redinbaugh M.G., Wadsworth G.J., Scandalios J.G.:
RL "Characterization of catalase transcripts and their differential
expression in maize."
CC Biochim. Biophys. Acta 951:104-116(1988).
CC -!- FUNCTION: DECOMPOSES HYDROGEN PEROXIDE IN WATER AND OXYGEN: SERVES
TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN PEROXIDE. ITS
LEVELS ARE HIGHEST IN THE LIGHT PERIOD AND ARE LOWEST IN THE DARK
PERIOD. HENCE MAY BE IMPORTANT FOR SCAVENGING HYDROGEN PEROXIDE
AT NIGHT, THAN DURING THE DAY.
CC AT CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
CC -!- COFACTOR: HEME GROUP.
CC -!- SUBUNIT: HOMOTETRAMER.
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC -!- TISSUE SPECIFICITY: LEAF MESOPHYLL CELLS, PERICARP, SEEDLING ROOTS
AND THE COLPOPILE.
CC -!- SIMILARITY: BELONGS TO THE CATALASE FAMILY.
CC -!- CAUTION: THERE ARE TWO EMBL ENTRIES THAT CORRESPOND TO THE REF.2
SEQUENCE AND THEY ENCODE SLIGHTLY DIFFERENT PROTEINS (SEE THE
FEATURE TABLE).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L05934; AAC37357.1;
DR EMBL; M33103; AAA33441.1;
DR EMBL; X12539; CAA31057.1;
DR PIR; JA0091; JAO091.
DR HSSP; P00432; 7CAT.
DR MA1ZEDB; 13855;
DR INTERPRO; IPR002226;
DR PFAM; PF00199; catalase; 1.
DR PRINTS; PR00067; CATALASE.
DR PROSITE; PS00437; CATALASE_1; 1.
DR PROSITE; PS00438; CATALASE_2; 1.
DR Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
KW Mitochondrion; Multigene family.
FT ACT_SITE 67 BY SIMILARITY.
FT ACT_SITE 140 140 PROXIMAL HEME LIGAND (BY SIMILARITY).
FT BINDING 351 351 A -> D (IN REF. 2).
FT CONFLICT 57 57 C -> S (IN REF. 2).
FT CONFLICT 79 79 T -> Q (IN REF. 2).
FT CONFLICT 107 107 H -> P (IN REF. 2).
FT CONFLICT 110 110 R -> P (IN REF. 2).
FT CONFLICT 112 112 PE -> GR (IN REF. 2).
FT CONFLICT 115 116 P -> A (IN REF. 2).
FT CONFLICT 121 121 D -> H (IN REF. 2).
FT CONFLICT 194 194 V -> E (IN REF. 2).
FT CONFLICT 234 236 C -> S (IN REF. 2).
FT CONFLICT 244 245 AL -> R (IN REF. 2).
FT CONFLICT 254 254 MISSING (IN REF. 2).
FT CONFLICT 264 264 A -> AE (IN REF. 2).
FT CONFLICT 280 281 DT -> AO (IN REF. 2).
FT CONFLICT 282 282 E -> O (IN REF. 2).
FT CONFLICT 282 282 F -> L (IN REF. 2).
FT CONFLICT 319 319 AH -> GT (IN REF. 2).
FT CONFLICT 374 375 F -> L (IN REF. 2).
FT CONFLICT 386 386 PLQAAP -> RRCGRAA (IN REF. 2).
FT CONFLICT 402 408 RRFADSLGHPKVSQ -> KAIRLRATPQOP (IN
REF. 2).
FT CONFLICT 452 465

VSIL_TRYBB STANDARD; PRT; 471 AA.
ID P26326;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE VARIANT SURFACE GLYCOPROTEIN ILTAT 1.21 PRECURSOR (VSG).
OS Trypanosoma brucei brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN-ISOLATE MIAG 201;
RX MEDLINE=92046037; PubMed=1942032;
RA Carlington M., Miller N., Blum M.L., Roditi I., Willey D.C.,
Turner M.J.;
RT "Variant specific glycoprotein of Trypanosoma brucei consists of two
domains each having an independently conserved pattern of cysteine
residues."
RL J. Mol. Biol. 221:823-835(1991).
CC -!- FUNCTION: VSG FORMS A COAT ON THE SURFACE OF THE PARASITE. THE
TRYPANOSOME EVADES THE IMMUNE RESPONSE OF THE HOST BY EXPRESSING
A SERIES OF ANTIGENICALLY DISTINCT VSGS FROM AN ESTIMATED 1000
VSG GENES.
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.
CC A SOLUBLE FORM IS RELEASED FROM RUPTURED CELLS BY THE ACTION OF A
PI-PLC.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X56766; CAA40085.1; -
DR PIR; S18446; S18446.
KW Glycoprotein; Antigen; Trypanosomiasis; GPI-anchor; Membrane; Signal.
FT SIGNAL 1 21
FT CHAIN 22 454 VARIANT SURFACE GLYCOPROTEIN ILTAT 1.21.
FT PROPEP 455 471 HYDROPHOBIC, REMOVED DURING MATURATION.
FT CARBOHYD 64 64 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 405 405 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 450 450 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 454 454 GPI-ANCHOR (BY SIMILARITY).
FT LIPID 454 454
FT SSEQUENCE 471 AA; 50480 MW; 784f48E944EEA37C CRC64;
Query Match 61.8%; Score 34; DB 1; Length 471;
Best Local Similarity 66.7%; Pred. NO. 27;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 GSSSHDTLC 9
|| || ||
Db 213 GSIGHDMLC 221
RESULT 11
CAT3_MAIZE STANDARD; PRT; 496 AA.
ID CAT3_MAIZE
AC P18123;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE CATALASE ISOZYME 3 (EC 1.11.1.6).
DN CAT3.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=CV. W64A; TISSUE=SEEDLING LEAF;
RX MEDLINE=94003062; PubMed=8400123;

FT CONFLICT 478 478 C -> V (IN REF. 2).
SQ SEQUENCE 496 AA; 56796 MW; 566FFD05B3795B49 CRC64;

Query Match 61.8%; Score 34; DB 1; Length 496;
Best Local Similarity 77.8%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SSSHDTLCT 10
|||||
Db 11 SSSHDTTVT 19

RESULT 12
NFX1_HUMAN
ID NFX1_HUMAN STANDARD; PRT; 1104 AA.
AC Q12986;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE TRANSCRIPTIONAL REPRESSOR NF-X1.
GN NFX1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-95053707; PubMed-7964459;

RA Song Z., Krishna S., Thanos D., Strominger J.L., Ono S.J.;
RT "A novel cysteine-rich sequence-specific DNA-binding protein
RT interacts with the conserved X-box motif of the human major
RT histocompatibility complex class II genes via a repeated Cys-His
RT domain and functions as a transcriptional repressor.";
RL J. Exp. Med. 180:1763-1774(1994)
CC -!- FUNCTION: REPRESSOR OF HLA-DRA TRANSCRIPTION. BINDS TO THE X-BOX
CC MOTIF OF CLASS II MHC GENES. MAY PLAY AN IMPORTANT ROLE IN
CC REGULATING THE DURATION OF AN INFLAMMATORY RESPONSE BY LIMITING
CC THE PERIOD IN WHICH CLASS II MHC MOLECULES ARE INDUCED BY
CC INTERFERON-GAMMA.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- INDUCTION: BY INTERFERON-GAMMA.
CC -!- SIMILARITY: TO D.MELANOGASTER SHUTTLE CRAFT PROTEIN (STC) AND
CC YEAST YNL023C.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; U15306; AAA69517.1; -;
DR HSSP; P02876; 2WGC.
DR INTERPRO; IPR000967; -;
DR INTERPRO; IPR001374; -;
DR PFAM; PF01424; R3H; 1.
DR PFAM; PF01422; zf-NF-X1; 8.

KW Transcription regulation; Repressor; DNA-binding; Nuclear protein;
KW Repeat.
FT DOMAIN 424 876 7 X APPROXIMATE REPEATS, CYS-RICH.
FT REPEAT 424 463 1.
FT REPEAT 480 517 2.
FT REPEAT 541 580 3.
FT REPEAT 606 647 4.
FT REPEAT 695 734 5.
FT REPEAT 806 842 6.
FT REPEAT 836 876 7.
SQ SEQUENCE 1104 AA; 123149 MW; 36FEC4292F78130C CRC64;

Query Match 61.8%; Score 34; DB 1; Length 1104;
Best Local Similarity 66.7%; Pred. No. 66;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 1 GSSSHDTLC 9
|||||
Db 518 GSTSRDVLVC 526
RESULT 13
CFLA_MCV1
ID CFLA_MCV1 STANDARD; PRT; 241 AA.
AC Q98325; O11298;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE VIRAL CASP8 AND FADD-LIKE APOPTOSIS REGULATOR (V-CFLAR) (VIRAL
DE FLICE-INHIBITORY PROTEIN) (V-FLIP).
GN 159L OR H-H2.2.
OS Molluscum contagiosum virus subtype 1 (MCV1).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Molluscipoxvirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-96325459; PubMed-8670425;
RA Senkevich T.G., Bugert J.J., Sisler J.R., Koonin E.V., Darai G.,
RA Moss B.;
RT "Genome sequence of a human tumorigenic poxvirus: prediction of
RT specific host response-evasion genes.";
RL Science 273:813-816(1996).
RN [2]
RP SEQUENCE OF 91-241 FROM N.A.
RX MEDLINE-97352177; PubMed-9208457;
RA Moratilla M., Agromayor M., Nunez A., Funes J.M., Varas A.J.,
RA Lopez-Esteban J.L., Esteban M., Martin-Gallardo A.;
RT "A random DNA sequencing, computer-based approach for the generation
RT of a gene map of molluscum contagiosum virus.";
RL Virus Genes 14:73-80(1997).
RN [3]
RP FUNCTION.
RX MEDLINE-97242415; PubMed-9087414;
RA Thome M., Schneider P., Hofmann K., Fickenscher H., Meinl E.,
RA Neipel F., Mattmann C., Burns K., Bodmer J.-L., Schroeter M.,
RA Scaffidi C., Kramer P.H., Peter M.E., Tschoopp J.;
RT "Viral FLICE-inhibitory proteins (FLIPs) prevent apoptosis induced by
RT death receptors.";
RL Nature 385:517-521(1997).
RN [4]
RP FUNCTION.
RX MEDLINE-97188440; PubMed-9037025;
RA Bertin J., Armstrong R.C., Ohtsuka S., Martin D.A., Wang Y., Banks S.,
RA Wang G.-H., Senkevich T.G., Alnemri E.S., Moss B., Lenardo M.J.,
RA Tomaselli K.J., Cohen J.I.;
RT "Death effector domain-containing herpesvirus and poxvirus proteins
RT inhibit both Fas- and TNFRI-induced apoptosis.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:1172-1176(1997).
RN [5]
RP CHARACTERIZATION.
RX MEDLINE-99457304; PubMed-10526240;
RA Tsukumo S.I., Yonehara S.;
RT "Requirement of cooperative functions of two repeated death effector
RT domains in caspase-8 and in MCL159 for induction and inhibition of
RT apoptosis, respectively.";
RL Genes Cells 4:541-549(1999).
CC -!- FUNCTION: INHIBITS TNFRSF1A, TNFRSF6 AND TNFRSF12 INDUCED
CC APOPTOSIS. MAY INTERFERE WITH CASPASE-8 RECRUITMENT AND ACTIVATION
CC AT THE DEATH-INDUCING SIGNALING COMPLEX (DISC). MAY LEAD TO HIGHER
CC VIRUS PRODUCTION AND CONTRIBUTE TO VIRUS PERSISTENCE AND
CC ONCOGENICITY.
CC -!- SUBUNIT: ASSOCIATES WITH THE DEATH-INDUCING SIGNALING COMPLEX
CC (DISC) FORMED BY TNFRSF6, FADD AND CASPASE-8. INTERACTS WITH FADD.
CC -!- SIMILARITY: CONTAINS 2 DEATH EFFECTOR DOMAINS (DED).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

Wed Jan 31 13:44:01 2001

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; U60315; AAC55287.1; -;
 DR EMBL; U86888; AAB57923.1; -;
 DR INTERPRO; IPR001875; -;
 DR PFAM; PF01335; DED; 2;
 DR PROSITE; PS0168; DED; 2;
 KW Apoptosis. 8 78 DED 1;
 FT DOMAIN 95 175 DED 2;
 FT DOMAIN 241 AA; 26939 MW; 155C9FB0B969E216 CRC64;
 SQ SEQUENCE

Query Match 60.0%; Score 33; DB 1; Length 241;
 Best Local Similarity 55.6%; Pred. No. 21;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GSSSHDTLTC 9

Db 204 GASEHEQLC 212

RESULT 14

ID CD40_MOUSE STANDARD; PRT; 289 AA.
 AC P27512;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE CD40L RECEPTOR PRECURSOR (B-CELL SURFACE ANTIGEN CD40) (BP50) (CDW40).
 GN TNFRSF5 OR CD40.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 [1]
 RN MEDLINE=92105763; PubMed=1370315;
 RX Torres R.M., Clark E.A.;
 RA "Differential increase of an alternatively polyadenylated mRNA
 RT species of murine CD40 upon B lymphocyte activation."
 RL J. Immunol. 148:620-626(1992).
 RN [2]
 RP REVISIONS.

RC STRAIN=BALB/C;
 RA Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 RL [3]

RN SEQUENCE FROM N.A.
 RC STRAIN=BALB/C; TISSUE=LIVER;
 RX MEDLINE=93094586; PubMed=1281194;
 RA Grimaldi J.C., Torres R., Kozak C.A., Chang R., Clark E.A.,
 RA Howard M., Cockayne D.A.;
 RT "Genomic structure and chromosomal mapping of the murine CD40 gene."
 RL J. Immunol. 149:3921-3926(1992).
 CC -1- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD40L.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- SIMILARITY: CONTAINS A LA-NGR/TNFR-TYPE CYSTEINE-RICH REGION.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; M83312; AAB08705.1; -;
 DR EMBL; M94126; AAA37404.1; -;
 DR EMBL; M94129; AAA37404.1; JOINED.

DR EMBL; M94128; AAA37404.1; JOINED.
 DR EMBL; M94127; AAA37404.1; JOINED.
 DR PIR; A46476; A46476.
 DR HSP; P25942; ICDF.
 DR MGI; 88336; TNFRSF5.
 DR INTERPRO; IPR001368; -;
 DR PFAM; PF00020; TNFR_C6; 4;
 DR PROSITE; PS00652; TNFR_NGFR_1; 1;
 DR PROSITE; PS00652; TNFR_NGFR_2; 4;
 KW Receptor; B-cell; Glycoprotein; Transmembrane; Repeat; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 289
 FT DOMAIN 20 193
 FT TRANSMEM 194 215
 FT DOMAIN 216 289
 FT DOMAIN 25 187
 FT REPEAT 25 60
 FT REPEAT 61 103
 FT REPEAT 104 144
 FT REPEAT 145 187
 FT CARBOHYD 153 153
 FT SEQUENCE 289 AA; 32111 MW; C791CB6D2FEA574E CRC64;
 SQ SEQUENCE

Query Match 60.0%; Score 33; DB 1; Length 289;
 Best Local Similarity 50.0%; Pred. No. 25;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GSSSHDTLCT 10

Db 95 GTSDDTCT 104

RESULT 15

YK95_MYCTU STANDARD; PRT; 316 AA.
 ID YK95_MYCTU
 AC Q10704;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE HYPOTHETICAL 33.8 KDA PROTEIN RV2095C PRECURSOR.
 GN RV2095C OR MTCY49.35C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacterium.
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., Mclean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter K., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
 RA Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 CC -1- SIMILARITY: STRONG, TO M.LEPRAE B2126_C2_220.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; Z73966; CAA98207.1; -;
 DR TUBERCULIST; RV2095C; -;
 KW Hypothetical protein; Signal.

FT SIGNAL 1 29 POTENTIAL.
FT CHAIN 30 316 HYPOTHETICAL PROTEIN RV2095C.
FT DOMAIN 119 141 ALA-RICH.
SQ SEQUENCE 316 AA; 33764 MW; 88D4ED2B9F5BAD76 CRC64;

Query Match 60.0%; Score 33; DB 1; Length 316;
Best Local Similarity 66.7%; Pred. No. 28;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SSSHDTLCT 10
:::|||||
Db 158 AASHDTLTT 166

Search completed: January 30, 2001, 17:05:59
Job time: 1220 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:54:43 ; Search time 254.27 seconds
(without alignments)
4.610 Million cell updates/sec

Title: US-09-518-931-2_COPY_185_194

Perfect score: 55

Sequence: 1 GSSSHDTLCT 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_15:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phase:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	100.0	300	4	095407
2	38	69.1	400	10	Q9M347
3	38	69.1	638	5	Q9U2B2
4	37	67.3	346	10	Q24513
5	37	67.3	346	10	Q9LPV9
6	37	67.3	523	5	Q77238
7	37	67.3	554	5	Q9VTX7
8	37	67.3	1157	5	Q9VP22
9	37	67.3	10797	3	Q94116
10	36	65.5	144	2	Q9KVY6
11	36	65.5	321	5	Q9VR53
12	36	65.5	337	10	Q24514
13	36	65.5	342	4	Q15491
14	36	65.5	342	10	Q9M610
15	36	65.5	346	10	Q38960
16	36	65.5	478	5	Q93759
17	36	65.5	611	5	Q19115
18	36	65.5	638	5	Q77036
19	36	65.5	1038	3	Q9P6L9

20	35	63.6	86	12	Q96742
21	35	63.6	114	12	Q96736
22	35	63.6	283	4	Q92956
23	35	63.6	283	4	Q9UM65
24	35	63.6	283	6	Q9XS28
25	35	63.6	372	4	Q9UHP4
26	35	63.6	401	4	Q00300
27	35	63.6	419	4	Q14722
28	35	63.6	481	10	Q81498
29	35	63.6	1463	5	Q9NKP3
30	34	61.8	119	12	Q89166
31	34	61.8	401	11	Q08712
32	34	61.8	401	11	Q08727
33	34	61.8	707	2	P73608
34	34	61.8	955	5	Q9W363
35	34	61.8	1023	11	P70588
36	34	61.8	1114	11	Q9JKW7
37	34	61.8	1243	5	Q9NGT8
38	33	60.0	147	2	Q9L6A2
39	33	60.0	151	12	O11298
40	33	60.0	163	10	Q9LTK7
41	33	60.0	225	5	P91241
42	33	60.0	241	12	Q98325
43	33	60.0	243	10	Q9L2H3
44	33	60.0	267	5	O61076
45	33	60.0	267	10	Q9SMP1

ALIGNMENTS

RESULT	1
ID	095407
AC	095407; PRELIMINARY; PRT: 300 AA.
DT	01-MAY-1999 (TREMBLrel. 10, Created)
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE	DECOY RECEPTOR 3 (M68) (M68C) (M68E).
GN	DCR3 OR TR6.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=99087326; PubMed=9872321;
RA	Pitti R.M., Marsters S.A., Lawrence D.A., Roy M., Kischkel F.C.,
RA	Dowd P., Huang A., Donahue C.J., Sherwood S.W., Baldwin D.T.,
RA	Godowski P.J., Wood W.I., Gurney A.L., Hillan K.J., Cohen R.L.,
RA	Goddard A.D., Botstein D., Ashkenazi A.
RT	"Genomic amplification of a decoy receptor for Fas ligand in lung and
RT	colon cancer.";
RL	Nature 396:699-703(1998).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	TISSUE=BLOOD;
RX	MEDLINE=99253915; PubMed=10318773;
RA	Yu K.Y., Kwon B., Ni J., Zhai Y., Ebner R., Kwon B.S.;
RT	"A newly identified member of tumor necrosis factor receptor
RT	superfamily (TR6) suppresses LIGHT-mediated apoptosis.";
RL	J. Biol. Chem. 274:13733-13736(1999).
RN	[3]
RP	SEQUENCE FROM N.A.
RC	TISSUE=PANCREAS;
RX	MEDLINE=20122600; PubMed=10655513;
RA	Bai C., Connolly B., Metzker M.L., Hilliard C.A., Liu X., Sandig V.,
RA	Soderman A., Galloway S.M., Liu Q., Austin C.P., Caskey C.T.;
RT	"Overexpression of M68/DCR3 in human gastrointestinal tract tumors
RT	independent of gene amplification and its location in a four-gene
RT	cluster.";
RL	Proc. Natl. Acad. Sci. U.S.A. 97:1230-1235(2000).
DR	EMBL; AF104419; AAD03056.1; -.

Wed Jan 31 13:44:02 2001

```

DR EMBL: AF134240; RAD29688.1; -.
DR EMBL: AF217796; RAD35244.1; -.
DR EMBL: AF217793; AAF33685.1; -.
DR EMBL: AF217794; AAF33686.1; -.
DR HSSP: P25942; ICDF.
DR INTERPRO: IPR000561; -.
DR INTERPRO: IPR001368; -.
DR PFAM: PF00020; TNFR_C6; 4.
DR PROSITE: PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
DR PROSITE: PS00050; TNFR_NGFR_2; 2.
DR PRODOM: PD000771; -. 1.
KW Receptor.
SQ SEQUENCE 300 AA; 32679 MW; F90AEE33718449AF CRC64;

Query Match 100.0%; Score 55; DB 4; Length 300;
Best Local Similarity 100.0%; Pred. No. 0.0023;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSSSHDTLCT 10
| | | | | | | |
Db 185 GSSSHDTLCT 194

RESULT 2
Q9M347 PRELIMINARY; PRT; 400 AA.
AC Q9M347;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE HYPOTHETICAL 45.0 KDA PROTEIN.
GN F5K20.90.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Monfort A., Casacuberta E., Puigdomenech P., Mewes H.W., Rudd S.,
RA Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;
RA Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RL [2]
RW SEQUENCE FROM N.A.
RP EU Arabidopsis sequencing project;
RP Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL: AL132960; CAB88341.1; -.
DR Hypothetical protein.
KW SEQUENCE 400 AA; 45032 MW; 624DF554C1E9951D CRC64;

Query Match 69.1%; Score 38; DB 10; Length 400;
Best Local Similarity 77.8%; Pred. No. 7.9;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SSSHDTLCT 10
| | | | |
Db 164 SSSHDTLCT 172

RESULT 3
Q9U2B2 PRELIMINARY; PRT; 638 AA.
AC Q9U2B2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Y47H10A.1 PROTEIN.
GN Y47H10A.1
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McLay K.;
RA Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RL [2]
RW SEQUENCE FROM N.A.
RA none;
RA "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL: AL132848; CAB60389.1; -.
DR HSSP: P17655; IDKV.
DR INTERPRO: IPR001300; -.
DR PFAM: PF00648; Peptidase_C2; 1.
DR PFAM: PF01067; Calpain_III; 1.
DR PRINTS: PR00704; CALPAIN.
SQ SEQUENCE 638 AA; 72754 MW; BA5329030D1D7D9A CRC64;

Query Match 69.1%; Score 38; DB 5; Length 638;
Best Local Similarity 60.0%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GSSSHDTLCT 10
| | | | |
Db 488 GCDHDTLCT 497

RESULT 4
Q24513 PRELIMINARY; PRT; 346 AA.
AC Q24513;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE ATAN11.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=COLUMBIA;
RA de Vetten N., Quattrocchio F., Mol J., Koes R.;
RA Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RL EMBL: U94746; AAC18912.1; -.
DR MENDEL; 27016; Arath; 2266; 27016.
DR INTERPRO: IPR001680; -.
DR PFAM: PF00400; WD40; 3.
DR PRINTS: PR00320; GPROTEINRPT.
DR PROSITE: PS00678; WD_REPEATS; UNKNOWN_2.
SQ SEQUENCE 346 AA; 39089 MW; AF61E5238E8265FE CRC64;

Query Match 67.3%; Score 37; DB 10; Length 346;
Best Local Similarity 70.0%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSSSHDTLCT 10
| | | | |
Db 152 GSSSHDTLCT 161

RESULT 5
Q9LPV9 PRELIMINARY; PRT; 346 AA.
AC Q9LPV9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)

```

```

DE F13K23.16 PROTEIN.
GN F13K23.16.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Liu S.X., Sakano H., Yu G., Lee J.M., Lenz C., Pham P., Toriumi M.,
RA Chin C., Chioi J., Choi E., Chung M., Gonzalez A., Howng B., Liu A.,
RA Vaysberg M., Altafi H., Brooks S., Buehler E., Chao Q., Conn L.,
RA Conway A.B., Hansen N.F., Johnson-Hopson C., Khan S., Kim C., Lam B.,
RA Miranda M., Nguyen M., Palm C.J., Shinn P., Southwick A., Davis R.W.,
RA Ecker J.R., Federspiel N.A., Theologis A.;
RT "The sequence of BAC F13K23 from Arabidopsis thaliana chromosome 1.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC012187; AAF78495.1; -
SQ SEQUENCE 346 AA; 39088 MW; 5F6BE5243854BE93 CRC64;

Query Match 67.3%; Score 37; DB 10; Length 346;
Best Local Similarity 70.0%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GSSSHDTLCT 10
Db 152 GSSSHDTTCT 161
|:||| ||||
|:||| ||||

RESULT 6
ID 077238 PRELIMINARY; PRT; 523 AA.
AC 077238;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE EYEGONE.
GN EYG.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Jun S., Wallen R.V., Goriely A., Kalionis B., Desplan C.;
RT "Lune/Eye gene, a new Pax-like protein, uses a partial paired domain
RT and a homeodomain for DNA recognition.";
RL Proc. Natl. Acad. Sci. U.S.A. 0:0-0(1998).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AF091739; AAC78330.1; -
DR HSSP; P06601; 1PDN.
DR INTERPRO; IPR001356; -
DR INTERPRO; IPR001523; -
DR PFAM; PF00046; homeobox; 1.
DR PFAM; PF00292; PAX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Nuclear protein.
SQ SEQUENCE 523 AA; 56495 MW; 9A33F12E86D9F67E CRC64;

Query Match 67.3%; Score 37; DB 5; Length 523;
Best Local Similarity 87.5%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GSSSHDTL 8
Db 203 GSSSHDTL 210
|:| |||||
|:| |||||

```

```

RESULT 7
Q9VTX7 PRELIMINARY; PRT; 554 AA.
ID 09VTX7;
AC 09VTX7;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE EYG PROTEIN.
GN EYG.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balleg R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhattacharya D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003541; AAF49918.1; -
DR HSSP; P06601; 1EJL.
DR FLYBASE; FBgn0000625; eyg.
DR INTERPRO; IPR001356; -
DR INTERPRO; IPR001523; -
DR PFAM; PF00046; homeobox; 1.
DR PFAM; PF00292; PAX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
SQ SEQUENCE 554 AA; 59552 MW; 4BAA8D65FCEAE114 CRC64;

```

```

Query Match 67.3%; Score 37; DB 5; Length 554;
Best Local Similarity 87.5%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GSSSHDTL 8

```

Best Local Similarity 100.0%; Pred. No. 34;		Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 GSSSHDT 7		
Db	340 GSSSHDT 346		
RESULT	9		
Q94116			
ID	O94116	PRELIMINARY;	PRT: 10797 AA.
AC	O94116;		
DT	01-MAY-1999 (TrEMBLrel. 10, Created)		
DT	01-MAY-1999 (TrEMBLrel. 10, Last sequence update)		
DT	01-OCT-2000 (TrEMBLrel. 15, Last annotation update)		
DE	PEPTIDE SYNTHETASE.		
OS	Aureobasidium pullulans.		
OC	Eukaryota; Fungi; Ascomycota; Ascomycota incertae sedis; Dothioraceae;		
OC	anamorphic Dothioraceae; Aureobasidium.		
OX	NCBI_TaxID=5580;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=RI06;		
RA	Peery R.B., Thornewell S.J., Tobin M.B., Skatrud P.L.;		
RT	"Discovery of an MDR-like gene adjacent to a peptide synthetase in		
RT	Aureobasidium pullulans."		
RL	Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.		
DR	EMBL: U85909; AAD00581.1; -		
DR	HSSP: P14687; IAMU.		
DR	INTERPRO: IPR000255; -		
DR	INTERPRO: IPR000873; -		
DR	INTERPRO: IPR001242; -		
DR	INTERPRO: IPR001680; -		
DR	INTERPRO: IPR001993; -		
DR	PFAM: PF00501; AMP-binding; 9.		
DR	PFAM: PF00550; PP-binding; 12.		
DR	PFAM: PF00668; DUF4; 11.		
DR	PROSITE: PS00012; PHOSPHOTANTHETINE; UNKNOWN_10.		
DR	PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_1.		
DR	PROSITE: PS00455; AMP_BINDING; 4.		
DR	PROSITE: PS00678; WD_REPEATS; UNKNOWN_1.		
DR	PROSITE: PS50075; ACP_DOMAIN; 12.		
SQ	SEQUENCE 10797 AA; 1190733 MW; 9EDB1A3E0A13A2D0 CRC64;		
Query Match 67.3%; Score 37; DB 3; Length 10797;			
Best Local Similarity 75.0%; Pred. No. 2.7e+02;			
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;			
QY	3 SSDDTICT 10		
Db	9444 ASHDLVCT 9451		
RESULT	10		
Q9KVI6			
ID	Q9KVI6	PRELIMINARY;	PRT: 144 AA.
AC	Q9KVI6;		
DT	01-OCT-2000 (TrEMBLrel. 15, Created)		
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)		
DT	01-OCT-2000 (TrEMBLrel. 15, Last annotation update)		
DE	MIOC PROTEIN.		
GN	VC0002.		
OS	Vibrio cholerae.		
OC	Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.		
OX	NCBI_TaxID=666;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=EL TOR N16961 / SEROTYPE O1;		
RC	MEDLINE=20406833; PubMed=10952301;		
RA	Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,		
RA	Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,		
RA	Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,		

RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Uterback T., Fleishmann R.D., Niernan W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*
 RT *cholerae*.";
 RL Nature 406:477-483(2000).
 DR EMBL; AE004093; AAF93180.1; -.
 DR TIGR: VC0002; -.
 SQ SEQUENCE 144 AA; 15662 MW; DA54074971621116 CRC64;

Query Match 65.5%; Score 36; DB 2; Length 144;
 Best Local Similarity 66.7%; Pred. No. 7.7;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GSSSHDTLC 9
 | |||||
 DB 89 GDSYDTFC 97

RESULT 11
 Q9VR53 PRELIMINARY; PRT; 321 AA.
 ID Q9VR53;
 AC Q9VR53;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE CG14614 PROTEIN.
 GN CG14614.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003574; AAF50953.1; -.
 DR FLYBASE; FBgn0031186; CG14614.
 DR INTERPRO; IPR001680; -.
 DR PFAM; PF00400; WD40; 2.
 DR PROSITE; PS00678; WD REPEATS; 1.
 SQ SEQUENCE 321 AA; 35955 MW; 124DC0047E8E4BF1 CRC64;

Query Match 65.5%; Score 36; DB 5; Length 321;
 Best Local Similarity 70.0%; Pred. No. 16;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GSSSHDTLC 10
 | || || ||
 DB 116 GTSSIDTTCT 125

RESULT 12
 O24514 PRELIMINARY; PRT; 337 AA.
 ID O24514;
 AC O24514;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE AN1.
 GN AN1.
 OS Petunia hybrida (Petunia).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
 OC Solanales; Solanaceae; Petunia.
 OC NCBI_TaxID=4102;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV V26; TISSUE=FLOWER COROLLA;
 RA de Vetten N., Quattrocchio F., Mol J., Koes R.;
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U94748; AAC18914.1; -.
 DR MENDEL; 27017; Pstx; 2266; 27017.
 DR INTERPRO; IPR001680; -.
 DR PFAM; PF00400; WD40; 2.
 DR PRINTS; PR00320; GPROTEINBRPT.
 DR PROSITE; PS00678; WD REPEATS; UNKNOWN_2.
 SQ SEQUENCE 337 AA; 37857 MW; 7024CAEDF5FD109C CRC64;

Query Match 65.5%; Score 36; DB 10; Length 337;
 Best Local Similarity 70.0%; Pred. No. 17;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GSSSHDTLC 10
 | || || ||
 DB 145 GTSSIDTTCT 154

RESULT 13
 O15491 PRELIMINARY; PRT; 342 AA.
 ID O15491;
 AC O15491;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE HAN1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;

Wed Jan 31 13:44:02 2001

```

de Vetten N., Quattrocchio F., Mol J., Koes R.;
RA Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RL EMBL: U94747; AAC18913.1; -.
DR INTERPRO: IFR001680; -.
DR PFAM: PF00400; WD40; 2.
DR PRINTS: PR00320; GPROTEINBRPT.
DR PROSITE: PS00678; WD_REPEATS; UNKNOWN_1.
SQ SEQUENCE 342 AA; 38926 MW; 794CC69A45D0CC7C CRC64;

Query Match 65.5%; Score 36; DB 4; Length 342;
Best Local Similarity 70.0%; Pred. No. 17;
Matches 7; Conservative 1; Mismatches 0; Gaps 0;

QY 1 GSSSHDTLCT 10
   1:|||||
Db 137 GTSSIDTCT 146

RESULT 14
Q9M610 PRELIMINARY; PRT; 342 AA.
ID Q9M610;
AC Q9M610;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE TTGL-LIKE PROTEIN.
OS Malus domestica (Apple) (Malus sylvestris).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Rosales; Rosaceae; Malus.
OC NCBI_TaxID=3750;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. MCINTOSH; TISSUE=RIPENING FRUIT;
RA Casas-Mollano A., Destefano-Beltran L.;
RT "Characterization of a cDNA encoding a Ttgl-like protein (Accession
   No. AF220203) from apple fruit (PGR00-041).";
RL Plant Physiol. 122:1458-1458(2000).
DR EMBL: AF220203; AAF27919.1; -.
SQ SEQUENCE 342 AA; 38473 MW; A79F506A5D563014 CRC64;

Query Match 65.5%; Score 36; DB 10; Length 342;
Best Local Similarity 70.0%; Pred. No. 17;
Matches 7; Conservative 1; Mismatches 0; Gaps 0;

QY 1 GSSSHDTLCT 10
   1:|||||
Db 150 GTSSIDTCT 159

RESULT 15
Q38960 PRELIMINARY; PRT; 346 AA.
ID Q38960;
AC Q38960;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE HYPOTHETICAL 39.1 KDA PROTEIN (BETA-TRANSDUCIN LIKE PROTEIN).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Quigley F., Dao P., Mache R.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;

```

```

EX MEDLINE=97086699; PubMed=8932388;
RA Quigley F., Dao P., Cottet A., Mache R.;
RT "Sequence analysis of an 81 kb contig from Arabidopsis thaliana
   chromosome III.";
RL Nucleic Acids Res. 24:4313-4318(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
   features of the regions of 4,504,864 bp covered by sixty P1 and TAC
   clones.";
RL DNA Res. 7:131-135(2000).
DR EMBL: X98130; CAA66815.1; -.
DR EMBL: X97488; CAA66120.1; -.
DR EMBL: AB026648; BAB01729.1; -.
DR MENDEL: 13066; Arath; 2266; L3066.
DR INTERPRO: IPR001680; -.
DR PFAM: PF00400; WD40; 3.
DR PRINTS: PR00320; GPROTEINBRPT.
DR PROSITE: PS00678; WD_REPEATS; UNKNOWN_2.
KW Hypothetical protein.
SQ SEQUENCE 346 AA; 39091 MW; C71DB62672451214 CRC64;

Query Match 65.5%; Score 36; DB 10; Length 346;
Best Local Similarity 70.0%; Pred. No. 17;
Matches 7; Conservative 1; Mismatches 0; Gaps 0;

QY 1 GSSSHDTLCT 10
   1:|||||
Db 152 GTSSIDTCT 161

Search completed: January 30, 2001, 16:54:46
Job time: 902 sec

```


THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:47:42 ; Search time 132.77 Seconds
(without alignments)
1.352 Million cell updates/sec

Title: US-09-518-931-2_COPY_185_194
Perfect score: 55
Sequence: 1 GSSSHDTLCT 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/PCrUS_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	55	100.0	300	2	US-08-794-796-2
2	36	65.5	162	2	US-08-219-237B-7
3	36	65.5	1245	1	US-08-158-232-8
4	36	65.5	1245	1	US-08-304-626-8
5	36	65.5	1245	2	US-08-611-928-8
6	36	65.5	1245	3	US-09-173-891-8
7	35	63.6	283	4	PCT-US96-12374-2
8	35	63.6	401	3	US-08-974-022-6
9	35	63.6	1205	1	US-07-908-245-2
10	35	63.6	1205	2	US-08-319-866-10
11	35	63.6	1205	3	US-09-123-708-6
12	35	63.6	1205	3	US-09-123-624-6
13	34	61.8	401	3	US-08-974-022-2
14	34	61.8	401	3	US-08-974-022-4
15	34	61.8	1104	2	US-08-327-832-5
16	34	61.8	1104	2	US-08-828-584-5
17	33	60.0	197	3	US-08-974-022-49
18	33	60.0	227	3	US-08-974-022-48
19	33	60.0	241	3	US-09-382-155-21
20	33	60.0	1455	2	US-08-726-012B-2
21	32	58.2	127	2	US-08-162-146-3
22	32	58.2	348	3	US-09-248-528-7
23	32	58.2	401	5	5252556-1
24	31	56.4	18	3	US-08-847-844A-50
25	31	56.4	39	1	US-08-050-319B-41
26	31	56.4	39	2	US-08-465-982-41
27	31	56.4	43	1	US-08-050-319B-34
28	31	56.4	43	2	US-08-465-982-34

29	31	56.4	103	3	US-08-946-329A-72	Sequence 72, Appl
30	31	56.4	163	2	US-08-219-237B-5	Sequence 5, Appl
31	31	56.4	164	2	US-08-232-087A-9	Sequence 9, Appl
32	31	56.4	311	2	US-08-602-359A-41	Sequence 41, Appl
33	31	56.4	461	1	US-08-385-229-2	Sequence 2, Appl
34	31	56.4	461	2	US-08-650-000-2	Sequence 2, Appl
35	31	56.4	461	5	5395760-2	Patent No. 5395760
36	31	56.4	486	1	US-08-243-010-1	Sequence 1, Appl
37	31	56.4	518	1	US-08-385-229-4	Sequence 4, Appl
38	31	56.4	543	3	US-09-199-229-2	Sequence 2, Appl
39	31	56.4	543	3	US-09-443-087-2	Sequence 2, Appl
40	31	56.4	610	1	US-08-410-804-1	Sequence 1, Appl
41	31	56.4	610	1	US-08-259-514-1	Sequence 1, Appl
42	31	56.4	610	2	US-08-858-311-1	Sequence 1, Appl
43	31	56.4	870	1	US-08-785-241-4	Sequence 4, Appl
44	31	56.4	875	1	US-08-785-241-5	Sequence 5, Appl
45	31	56.4	2465	2	US-08-596-291-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-794-796-2
; Sequence 2, Application US/08794796
; Patent No. 5885800
; GENERAL INFORMATION:
; APPLICANT: Emery, John
; APPLICANT: Tan, KB
; APPLICANT: Truneh, Alem
; APPLICANT: Young, Peter
; TITLE OF INVENTION: Tumor Necrosis Related Receptor,
; TITLE OF INVENTION: TR4
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithkline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/794,796
; FILING DATE: 04-FEB-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: GH50000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-4026
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 300 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-794-796-2

Query Match 100.0%; Score 55; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 0.012;

```

; TITLE OF INVENTION: No. 5596071el Bacillus thuringiensis Toxins Active
; TITLE OF INVENTION: Against Hymenopteran Pests
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/158,232
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/887,980
; FILING DATE: 22-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/797,645
; FILING DATE: 25-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/703,977
; FILING DATE: 22-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: M/SCJ104.C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1245 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: BACILLUS THURINGIENSIS
; INDIVIDUAL ISOLATE: PS86Q3
; IMMEDIATE SOURCE:
; LIBRARY: LAMBDA GEM (cm) - 11 library
; CLONE: 86Q3A
; US-08-158-232-8

Query Match 65.5%; Score 36; DB 1; Length 1245;
Best Local Similarity 77.8%; Pred. NO. 1.3e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GSSSHDTLCT 10
Db 773 SSEHDTLAT 781

RESULT 4
US-08-304-626-8
; Sequence 8, Application US/08304626
; Patent No. 5616495
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel M.
; APPLICANT: Kennedy, M. Keith
; APPLICANT: Randall, John Brooks
; APPLICANT: Meier, Henry
; APPLICANT: Focerrada, Luis
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Schwab, George E.
; APPLICANT: Fu, Jenny

; TITLE OF INVENTION: DNA Coding for Human Cell Surface Antigen
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James W. Hellwege
; STREET: P.O. Box 2366 Eads Station
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/219,237B
; FILING DATE: 28-MAR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/872,129
; FILING DATE: 22-APR-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: James W. Hellwege
; REGISTRATION NUMBER: 28,808
; REFERENCE/DOCKET NUMBER: 516762
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 162 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-219-237B-7

Query Match 65.5%; Score 36; DB 2; Length 162;
Best Local Similarity 60.0%; Pred. No. 15;
Matches 6; Conservative 2; Mismatches 2; Indels 2; Gaps 0;

QY 1 GSSSHDTLCT 10
Db 71 GTSETDTICT 80

RESULT 3
US-08-158-232-8
; Sequence 8, Application US/08158232
; Patent No. 5596071
; GENERAL INFORMATION:
; APPLICANT: Payne, Jewel
; APPLICANT: Kennedy, M. Keith
; APPLICANT: Randall, John Brooks
; APPLICANT: Meier, Henry
; APPLICANT: Focerrada, Luis
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Schwab, George E.
; APPLICANT: Fu, Jenny

```

APPLICANT: Schnepf, Harry E.
APPLICANT: Schwab, George E.
TITLE OF INVENTION: No. 5616495el Bacillus thuringiensis Isolates
TITLE OF INVENTION: Active Against Hymenopteran Pests and Genes Encoding
TITLE OF INVENTION: Hymenopteran-Active Toxins
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/304,626
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/887,980
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: M/SCJ 104
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1245 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: BACILLUS THURINGIENSIS
INDIVIDUAL ISOLATE: PS86Q3
IMMEDIATE SOURCE:
LIBRARY: LAMBDA GEM (tm) - 11 library
CLONE: 86Q3A
US-08-304-626-8

Query Match 65.5%; Score 36; DB 1; Length 1245;
Best Local Similarity 77.8%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 SSSHDTLCT 10
|| |||||
Db 773 SSEHDTLAT 781

RESULT 5
US-08-611-928-8
Sequence 8, Application US/08611928
Patent No. 5824792
GENERAL INFORMATION:
APPLICANT: Payne, Jewel
APPLICANT: Kennedy, M. Keith
APPLICANT: Randall, John Brooks
APPLICANT: Meier, Henry
APPLICANT: Ulick, Heidi Jane
APPLICANT: Foncecrada, Luis
APPLICANT: Schnepf, H. Ernest
APPLICANT: Schwab, George E.
APPLICANT: Fu, Jenny

TITLE OF INVENTION: No. 5824792el Bacillus thuringiensis Toxins Active
TITLE OF INVENTION: Against Hymenopteran Pests
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/611,928
FILING DATE: 06-MAR-1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/158,232
FILING DATE: 24-NOV-1993
APPLICATION NUMBER: US 07/887,980
FILING DATE: 22-MAY-1992
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/797,645
FILING DATE: 25-NOV-1991
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/703,977
FILING DATE: 22-MAY-1991
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: M/SCJ104.C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1245 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: BACILLUS THURINGIENSIS
INDIVIDUAL ISOLATE: PS86Q3
IMMEDIATE SOURCE:
LIBRARY: LAMBDA GEM (tm) - 11 library
CLONE: 86Q3A
US-08-611-928-8

Query Match 65.5%; Score 36; DB 2; Length 1245;
Best Local Similarity 77.8%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 SSSHDTLCT 10
|| |||||
Db 773 SSEHDTLAT 781

RESULT 6
US-09-173-891-8
Sequence 8, Application US/09173891
Patent No. 6077937
GENERAL INFORMATION:
APPLICANT: Payne, Jewel

RESULT 7

```

PCT-US96-12374-2
;
; Sequence 2, Application PC/TUS9612374
; GENERAL INFORMATION:
; APPLICANT: Northwestern University
; TITLE OF INVENTION: Herpes Virus Entry Mediator
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESSES:
; ADDRESSES: Dressler, Goldsmith, Milnamow & Katz, Ltd.
; STREET: 180 N. Stetson, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60601
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/12374
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Northrup, Thomas E.
; REGISTRATION NUMBER: 33,268
; REFERENCE/DOCKET NUMBER: NOR3446P020PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5400
; TELEFAX: (312) 616-5460
; TELEX: --
;
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 283 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US96-12374-2
;
; Query Match 63.6%; Score 35; DB 4; Length 2
; Best Local Similarity 66.7%; Pred. No. 41;
; Matches 6; Conservative 1; Mismatches 2; Indel
;
; QY 1 GSSSDHTLC 9
; 1: | | | | |
; Db 154 GTESQDTLC 162
;
; RESULT 8
; US-08-974-022-6
; Sequence 6, Application US/08974022
; Patent No. 6015938
; GENERAL INFORMATION:
; APPLICANT: Boyle, William J.
; APPLICANT: Lacey, David L.
; APPLICANT: Calzone, Frank J.
; APPLICANT: Chang, Ming-Shi
; TITLE OF INVENTION: OSTEOPROTEGERIN
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,022

```

;; FILING DATE: 12-DEC-1995
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/577,788
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Winter, Robert B.
;; REFERENCE/DOCKET NUMBER: A-378
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 401 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-974-022-6

Query Match 63.6%; Score 35; DB 3; Length 401;
Best Local Similarity 40.0%; Pred. NO. 60;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSSSHDTLCT 10
Db 177 GNATHONICS 186

RESULT 9

US-07-908-245-2
; Sequence 2, Application US/07908245
; Patent No. 5498539
; GENERAL INFORMATION:
; APPLICANT: Harrison, David G.
; APPLICANT: Alexander, R. Wayne
; APPLICANT: Murphy, T.J.
; APPLICANT: Nishida, Ken'ichi
; TITLE OF INVENTION: Endothelial Nitric Oxide Synthase
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/908,245
; FILING DATE: 19920702
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: EMU 111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6508
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1205 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Bovine
; TISSUE TYPE: Aorta

;; CELL TYPE: Endothelial
;; FEATURE:
; NAME/KEY: Binding-site
; LOCATION: 496..512
; OTHER INFORMATION: /note= "CA++/CAM binding domain"
;; FEATURE:
; NAME/KEY: Binding-site
; LOCATION: 651..678
; OTHER INFORMATION: /note= "FMN binding domain"
;; FEATURE:
; NAME/KEY: Binding-site
; LOCATION: 795..806
; OTHER INFORMATION: /note= "FAD-pyrophosphate binding
; OTHER INFORMATION: domain"
;; FEATURE:
; NAME/KEY: Binding-site
; LOCATION: 937..947
; OTHER INFORMATION: /note= "FAD-Isolalloxanthine
; OTHER INFORMATION: binding domain"
;; FEATURE:
; NAME/KEY: Binding-site
; LOCATION: 1012..1030
; OTHER INFORMATION: /note= "NADPH-Ribose binding
; OTHER INFORMATION: domain"
;; FEATURE:
; NAME/KEY: Binding-site
; LOCATION: 1111..1124
; OTHER INFORMATION: /note= "NADPH-Ribose binding
; OTHER INFORMATION: domain"
;; FEATURE:
; NAME/KEY: Domain
; LOCATION: 33..34
; OTHER INFORMATION: /note= "Potential proline directed
; OTHER INFORMATION: phosphorylation site"
;; FEATURE:
; NAME/KEY: Domain
; LOCATION: 46..47
; OTHER INFORMATION: /note= "Potential proline directed
; OTHER INFORMATION: phosphorylation site"
;; FEATURE:
; NAME/KEY: Domain
; LOCATION: 53..54
; OTHER INFORMATION: /note= "Potential proline directed
; OTHER INFORMATION: phosphorylation site"
;; FEATURE:
; NAME/KEY: Domain
; LOCATION: 58..59
; OTHER INFORMATION: /note= "Potential proline directed
; OTHER INFORMATION: phosphorylation site"
;; FEATURE:
; NAME/KEY: Domain
; LOCATION: 97..98
; OTHER INFORMATION: /note= "Potential proline directed
; OTHER INFORMATION: phosphorylation site"
;; FEATURE:
; NAME/KEY: Domain
; LOCATION: 116..117
; OTHER INFORMATION: /note= "Potential proline directed
; OTHER INFORMATION: phosphorylation site"
;; FEATURE:
; NAME/KEY: Domain
; LOCATION: 282..283
; OTHER INFORMATION: /note= "Potential proline
; OTHER INFORMATION: phosphorylation site"
;; FEATURE:
; NAME/KEY: Domain
; LOCATION: 459..460
; OTHER INFORMATION: /note= "Potential proline directed
; OTHER INFORMATION: phosphorylation site"
;; FEATURE:
; NAME/KEY: Domain
; LOCATION: 472..473
; OTHER INFORMATION: /note= "Potential proline directed

OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 602..603
OTHER INFORMATION: /note= "Potential proline directed
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 727..728
OTHER INFORMATION: /note= "Potential proline directed
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 838..839
OTHER INFORMATION: /note= "Potential proline directed
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 869..870
OTHER INFORMATION: /note= "Potential proline directed
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 872..873
OTHER INFORMATION: /note= "Potential proline directed
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 1085..1086
OTHER INFORMATION: /note= "Potential proline directed
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 1202..1203
OTHER INFORMATION: /note= "Potential proline directed
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 114..116
OTHER INFORMATION: /note= "cAMP dependent
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 141..143
OTHER INFORMATION: /note= "cAMP dependent
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 168..170
OTHER INFORMATION: /note= "cAMP dependent
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 633..635
OTHER INFORMATION: /note= "cAMP dependent
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 836..838
OTHER INFORMATION: /note= "cAMP dependent
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 1051..1053
OTHER INFORMATION: /note= "cAMP dependent
OTHER INFORMATION: phosphorylation site"
FEATURE:
NAME/KEY: Domain
LOCATION: 738..740
OTHER INFORMATION: /note= "cAMP dependent
OTHER INFORMATION: phosphorylation site"
US-07-908-245-2

Query Match 63.6%; Score 35; DB 1; Length 1205;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 GSSSHDTLC 9
||:||||
Db 79 GSITYDTLC 87
RESULT 10
US-08-319-866-10
; Sequence 10, Application US/08319866
; Patent No. 5929223
; GENERAL INFORMATION:
; APPLICANT: Tully, Timothy P.
; APPLICANT: Yin, Jerry C.
; APPLICANT: Regulski, Michael
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF GENES
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/319,866
; FILING DATE: 7-OCT-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: CSHL94-03
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1205 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-319-866-10

Query Match 63.6%; Score 35; DB 2; Length 1205;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 GSSSHDTLC 9
||:||||
Db 79 GSITYDTLC 87
RESULT 11
US-09-123-708-6
; Sequence 6, Application US/09123708
; Patent No. 6146887
; GENERAL INFORMATION:
; APPLICANT: SCHRADER, Juergen
; APPLICANT: GOEDECKE, Axel

```
/ TITLE OF INVENTION: DNA EXPRESSION VECTORS FOR USE IN GENE THERAPEUTIC
/ FILE OF INVENTION: TREATMENT OF VASCULAR DISORDERS
/ FILE REFERENCE: 511169-2003
/ CURRENT APPLICATION NUMBER: US/09/123,708
/ CURRENT FILING DATE: 1998-07-28
/ EARLIER APPLICATION NUMBER: 08/553,503
/ EARLIER FILING DATE: 1996-03-01
/ EARLIER APPLICATION NUMBER: P4411402.8
/ EARLIER FILING DATE: 1994-03-31
/ NUMBER OF SEQ ID NOS: 6
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 6
/ LENGTH: 1205
/ TYPE: PRT
/ ORGANISM: Cytomegalovirus
US-09-123-708-6

Query Match      63.6%; Score 35; DB 3; Length 1205;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSSSHDTLC 9
Db 79 GSITYDTLC 87

RESULT 12
US-09-123-624-6
/ Sequence 6, Application US/09123624
/ Patent No. 6149936
/ GENERAL INFORMATION:
/ APPLICANT: SCHRADER, Jurgen
/ APPLICANT: CODECKE, Axel
/ TITLE OF INVENTION: DNA EXPRESSION VECTORS FOR USE IN THE GENE THERAPEUTIC
/ FILE OF INVENTION: TREATMENT OF VASCULAR DISORDERS
/ FILE REFERENCE: 511169-2004
/ CURRENT APPLICATION NUMBER: US/09/123,624
/ CURRENT FILING DATE: 1998-07-28
/ PRIOR APPLICATION NUMBER: 08/553,503
/ PRIOR FILING DATE: 1996-03-01
/ PRIOR APPLICATION NUMBER: 4411402.8
/ PRIOR FILING DATE: 1994-03-31
/ NUMBER OF SEQ ID NOS: 6
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 6
/ LENGTH: 1205
/ TYPE: PRT
/ ORGANISM: Bos taurus
US-09-123-624-6

Query Match      63.6%; Score 35; DB 3; Length 1205;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSSSHDTLC 9
Db 79 GSITYDTLC 87

RESULT 13
US-08-974-022-2
/ Sequence 2, Application US/08974022
/ Patent No. 6015938
/ GENERAL INFORMATION:
/ APPLICANT: Boyle, William J.
/ APPLICANT: Lacey, David L.
/ APPLICANT: Calzone, Frank J.
/ APPLICANT: Chang, Ming-Shi
/ TITLE OF INVENTION: OSTEOPROTEGERIN
/ NUMBER OF SEQUENCES: 53
/ CORRESPONDENCE ADDRESS:
/
```

```
/ ADDRESSEE: Amgen Inc.
/ STREET: 1840 Behavilland Drive
/ CITY: Thousand Oaks
/ STATE: California
/ COUNTRY: USA
/ ZIP: 91320-1789
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/974,022
/ FILING DATE: 12-DEC-1995
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/577,788
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Winter, Robert B.
/ REFERENCE/DOCKET NUMBER: A-378
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 401 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-974-022-2

Query Match      61.8%; Score 34; DB 3; Length 401;
Best Local Similarity 40.0%; Pred. No. 90;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSSSHDTLC 10
Db 177 GNATHDNVCS 186

RESULT 14
US-08-974-022-4
/ Sequence 4, Application US/08974022
/ Patent No. 6015938
/ GENERAL INFORMATION:
/ APPLICANT: Boyle, William J.
/ APPLICANT: Lacey, David L.
/ APPLICANT: Calzone, Frank J.
/ APPLICANT: Chang, Ming-Shi
/ TITLE OF INVENTION: OSTEOPROTEGERIN
/ NUMBER OF SEQUENCES: 53
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Amgen Inc.
/ STREET: 1840 Behavilland Drive
/ CITY: Thousand Oaks
/ STATE: California
/ COUNTRY: USA
/ ZIP: 91320-1789
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/974,022
/ FILING DATE: 12-DEC-1995
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/577,788
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Winter, Robert B.
/ REFERENCE/DOCKET NUMBER: A-378
/ INFORMATION FOR SEQ ID NO: 4:
/
```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 401 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-974-022-4

Query Match 61.8%; Score 34; DB 3; Length 401;
Best Local Similarity 40.0%; Pred. No. 90;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSSSHDTLCT 10
I:::|:|:
Db 177 GNATHDNVCS 186

RESULT 15
US-08-327-832-5
; Sequence 5, Application US/08327832
; Patent No. 5840832
; GENERAL INFORMATION:
; APPLICANT: Ono, Santa J.
; APPLICANT: Strominger, Jack L.
; TITLE OF INVENTION: Transcription Factor Regulating MHC
; TITLE OF INVENTION: Expression, cDNA and Genomic Clones Encoding Same and
; TITLE OF INVENTION: Retroviral Expression Constructs Thereof
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner, Birch, McKie & Beckett
; STREET: 1001 G Street, N.W.
; CITY: Washington, D.C.
; STATE: District of Columbia
; COUNTRY: U.S.A.
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/327,832
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Posorske, Laurence H.
; REGISTRATION NUMBER: 34,698
; REFERENCE/DOCKET NUMBER: 1107.46362
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9153
; TELEFAX: 202-508-9299
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1104 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-327-832-5

Query Match 61.8%; Score 34; DB 2; Length 1104;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GSSSHDTLCT 9
I::|:|:|:
Db 518 GSTSRDVLCT 526

Search completed: January 30, 2001, 16:47:43
Job time: 619 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:45:23 ; Search time 183.67 seconds

(without alignments)
2.420 Million cell updates/sec

Title: US-09-518-931-2_COPY_205_217

Perfect score: 68

Sequence: 1 VPGAECERAVID 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq_36:*
- 1: /SIDSI1/gcgdata/geneseq/geneseq/AA1980.DAT:*
 - 2: /SIDSI1/gcgdata/geneseq/geneseq/AA1981.DAT:*
 - 3: /SIDSI1/gcgdata/geneseq/geneseq/AA1982.DAT:*
 - 4: /SIDSI1/gcgdata/geneseq/geneseq/AA1983.DAT:*
 - 5: /SIDSI1/gcgdata/geneseq/geneseq/AA1984.DAT:*
 - 6: /SIDSI1/gcgdata/geneseq/geneseq/AA1985.DAT:*
 - 7: /SIDSI1/gcgdata/geneseq/geneseq/AA1986.DAT:*
 - 8: /SIDSI1/gcgdata/geneseq/geneseq/AA1987.DAT:*
 - 9: /SIDSI1/gcgdata/geneseq/geneseq/AA1988.DAT:*
 - 10: /SIDSI1/gcgdata/geneseq/geneseq/AA1989.DAT:*
 - 11: /SIDSI1/gcgdata/geneseq/geneseq/AA1990.DAT:*
 - 12: /SIDSI1/gcgdata/geneseq/geneseq/AA1991.DAT:*
 - 13: /SIDSI1/gcgdata/geneseq/geneseq/AA1992.DAT:*
 - 14: /SIDSI1/gcgdata/geneseq/geneseq/AA1993.DAT:*
 - 15: /SIDSI1/gcgdata/geneseq/geneseq/AA1994.DAT:*
 - 16: /SIDSI1/gcgdata/geneseq/geneseq/AA1995.DAT:*
 - 17: /SIDSI1/gcgdata/geneseq/geneseq/AA1996.DAT:*
 - 18: /SIDSI1/gcgdata/geneseq/geneseq/AA1997.DAT:*
 - 19: /SIDSI1/gcgdata/geneseq/geneseq/AA1998.DAT:*
 - 20: /SIDSI1/gcgdata/geneseq/geneseq/AA1999.DAT:*
 - 21: /SIDSI1/gcgdata/geneseq/geneseq/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	100.0	153	20	Y22222
2	68	100.0	215	20	W93585
3	68	100.0	245	20	Y28449
4	68	100.0	271	20	Y42184
5	68	100.0	273	20	Y42185
6	68	100.0	300	19	W66102
7	68	100.0	300	19	W63622
8	68	100.0	300	20	Y03099
9	68	100.0	300	20	Y42182
10	68	100.0	300	20	Y17479
11	68	100.0	300	20	Y06817
12	68	100.0	300	20	W97749

13	68	100.0	300	20	W95082
14	68	100.0	300	21	Y77458
15	68	100.0	302	20	Y42183
16	45	66.2	23	12	R14721
17	45	66.2	437	15	R49740
18	45	66.2	437	16	R77840
19	45	66.2	437	17	W04428
20	38	55.9	531	21	Y84361
21	37	54.4	229	19	W52985
22	37	54.4	310	20	Y21590
23	37	54.4	429	21	Y86442
24	37	54.4	470	21	Y86441
25	37	54.4	687	21	Y78981
26	35	51.5	366	20	Y48456
27	35	51.5	384	19	W98586
28	35	51.5	467	20	Y74073
29	35	51.5	3396	14	R43662
30	34.5	50.7	112	20	Y60194
31	34	50.0	99	21	Y86548
32	34	50.0	134	19	Y85882
33	34	50.0	220	21	Y86551
34	34	50.0	278	21	Y53886
35	34	50.0	311	17	R96108
36	34	50.0	319	17	R96534
37	34	50.0	395	6	P50322
38	34	50.0	558	20	Y25911
39	34	50.0	558	20	Y25914
40	34	50.0	604	21	Y57087
41	34	50.0	985	20	W89419
42	34	50.0	985	20	W89415
43	34	50.0	1000	20	W89414
44	34	50.0	1000	20	W89418
45	34	50.0	3224	19	W54235

ALIGNMENTS

RESULT 1

Y22222
ID Y22222 standard; Protein; 153 AA.

AC Y22222;

DT 16-SEP-1999 (first entry)

DE Human TNFR superfamily soluble receptor protein sequence.

XX TNFLL; human; TNFR superfamily; tumour necrosis factor ligand; TNF;
KW tumour necrosis factor receptor; TNFR superfamily; cell proliferation;
KW cell differentiation; cytokine production; immunoglobulin; hyperplasia;
KW apoptosis inducer; activated T cell; autoimmune disease; inhibitor;
KW myasthenia gravis; insulin-dependent diabetes mellitus; endotoxin shock;
KW rheumatoid arthritis; multiple sclerosis; systemic lupus erythematosus;
KW tumour; proliferative disorder; neoplasia; dysplasia; immunocompetence;
KW lymphoid organogenesis; bacterial resistance; contact hypersensitivity;
KW delayed type sensitivity; therapy.

OS Homo sapiens.

PN W09933980-A2.

PD 08-JUL-1999.

PF 22-DEC-1998; 98WO-US27474.

PR 16-DEC-1998; 98US-0212270.

PR 30-DEC-1997; 97US-0068959.

PA (CHIR) CHIRON CORP.

PI Kassam A, Lamson G, Pot D, Tribouley C;

Wed Jan 31 13:44:08 2001

DR WPI; 1999-405508/34.
DR N-PSDB; X84621.
XX
XX New tumour necrosis factor ligands, useful for induction of cell
PT death and/or proliferation of cells
XX
XX
PS Claim 1; Page 61; 69pp; English.
XX
XX This sequence represents a tumour necrosis factor receptor (TNFR)
CC superfamily soluble protein of the invention. The invention also relates
CC to tumour necrosis factor (TNF) ligand (TNFL) family proteins. The TNFL
CC proteins play regulatory roles in cell proliferation and/or
CC differentiation, e.g. they can induce production of cytokines,
CC immunoglobulins, etc. A variety of diseases can be treated by modulating
CC the activity of TNFL proteins, e.g. they can induce apoptosis of
CC activated T cells but rescue resting T cell from apoptosis. TNFL
CC polypeptides can therefore be used to treat autoimmune diseases, such as
CC myasthenia gravis, insulin-dependent diabetes mellitus, rheumatoid
CC arthritis, multiple sclerosis, and systemic lupus erythematosus. TNFL
CC proteins also have tumour stimulating properties, so tumours can be
CC treated by inhibiting the expression or activity of TNFL. Other
CC proliferative disorders, such as neoplasias, dysplasias, and hyperplasia
CC can also be treated using TNFL inhibitors. The TNFL polypeptides and
CC polynucleotides can also be used to enhance or decrease TNF activity,
CC thus providing therapeutic benefits such as induction of cell death,
CC lymphoid organogenesis, or host bacterial resistance, and inhibition of
CC endotoxic shock, contact hypersensitivity, delayed type sensitivity or
CC immunocompetence of a transplant recipient. TNF and its receptors play a
CC major role in host defence and immunosurveillance. As such, there is a
CC need to identify new members of TNFR families. This invention provides
CC this need.
XX
XX Sequence 153 AA;
SQ

Query Match 100.0%; Score 68; DB 20; Length 153;
Best Local Similarity 100.0%; Pred. No. 7.5e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VPGAEECERAVID 13
Db 121 vpgaeceravid 132
|||||
RESULT 2
ID W93585 standard; Protein: 215 AA.
XX W93585;
XX
DT 18-JUN-1999 (first entry)
XX
DE Human hAPO6 protein.
XX
XX Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
KW developmental abnormality; gestational abnormality; prostate cancer;
KW APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
KW apoptosis; human.
XX
XX Homo sapiens.
XX
XX WO9911791-A2.
XX
XX
PD 11-MAR-1999.
XX
XX
PF 04-SEP-1998; 98WO-US18393.
XX
XX
PR 05-SEP-1997; 97US-0924634.
XX
XX (UNIW) UNIV WASHINGTON.
XX
XX Chaudhary PW;
PI

XX
DR WPI; 1999-205191/17.
DR N-PSDB; X23419.
XX
XX New Tumour Necrosis Factor family receptor polypeptides and ligands -
PT useful for diagnosis and treatment of prostate cancer and
PT developmental or gestational abnormalities
XX
XX Claim 29; Fig 9; 156pp; English.
PS
XX This invention describes isolated Tumour Necrosis Factor (TNF) family
CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
CC their active fragments. APO4 is useful for diagnosing prostate cancer
CC by determining levels of APO4 in an individual. Prostate cancer can also
CC be treated using APO4 selective binding agents linked to a therapeutic
CC moiety. APO4 polypeptides are also useful for identifying selective
CC binding agents, useful in diagnosis/treatment of disease by binding of
CC agents to the polypeptide/active fragment which is extracellular, or
CC expressed on the cell surface. The binding is preferably performed in
CC vivo. APO4 polypeptides/active fragments are also useful for screening
CC for agonists and antagonists by binding and observing the change in APO4
CC activity. Effective pharmacological agents useful in diagnosis or
CC treatment of disease are also identified using APO4 polypeptides/active
CC fragments and APO4 signal transducer molecules that specifically interact
CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4
CC activity. The method is performed in vivo or in vitro. APO polypeptides
CC are all useful as immunogens for preparing antibodies. APO4 is also
CC useful for diagnosis/treatment of developmental or gestational
CC abnormalities. APO8 was transfected to human breast carcinoma cell line
CC MCF-7, and induced apoptosis.
XX
XX Sequence 215 AA;
SQ

Query Match 100.0%; Score 68; DB 20; Length 215;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VPGAEECERAVID 13
Db 120 vpgaeceravid 132
|||||
RESULT 3
ID Y28449 standard; Protein: 245 AA.
XX Y28449;
XX
DT 29-SEP-1999 (first entry)
XX
DE A human tumour necrosis factor-R2-like proteins (TR2P)-1.
XX
XX Human tumour necrosis factor-R2-like protein; TR2P; achondroplasia;
KW osteoporosis; developmental disorder; Cushing's syndrome;
KW muscular dystrophy; epilepsy; hereditary neuropathy;
KW Charcot-Marie-Tooth disease; neurofibromatosis; hypothyroidism;
KW hydrocephalus; seizure disorder; cerebral palsy; spinal bifida;
KW congenital glaucoma; cataract; sensorineural hearing loss;
KW reproductive disorder; infertility; ovulatory defect; endometriosis;
KW autoimmune disorder; ectopic pregnancy; teratogenesis; spermatogenesis;
KW immunological disorder; AIDS; Addison's disease; allergy; bronchitis;
KW atherosclerosis; diabetes mellitus; Chron's disease; lupus;
KW irritable bowel syndrome; multiple sclerosis; infection;
KW neoplastic disorder; adenocarcinoma; leukaemia; lymphoma; melanoma;
KW myeloma; sarcoma.
XX
XX Homo sapiens.
XX
XX WO9931128-A2.
XX
XX 24-JUN-1999.
PD

XX PF 02-DEC-1998; 98WO-US25649.
XX PR 16-DEC-1997; 97US-0991945.
XX PA (INCY-) INCYTE PHARM INC.
XX PI Au-Young J, Bandman O, Hillman JL, Kaser MR, Tang YT;
XX PI WPI; 1999-457916/38.
XX DR N-PSDB; X89503.
XX
XX New tumour necrosis factor-R2-like protein - useful in the treatment
XX PT of osteogenesis, developmental, reproductive, immunological and
XX PT neoplastic disorders
XX PS Claim 1; Fig 1A-C; 81pp; English.
XX
XX The present sequence represents a human tumour necrosis factor-R2-like
XX CC protein (TR2P). The protein is used to treat and prevent osteogenesis,
XX CC developmental, reproductive, immunological and neoplastic disorders, and
XX CC also to diagnose disorders associated with TR2 protein expression. Such
XX CC disorders include osteogenesis disorders such as achondroplasia and
XX CC osteoporosis, developmental disorders such as Cushing's syndrome,
XX CC muscular dystrophy, and epilepsy, hereditary neuropathies such as
XX CC Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism,
XX CC hydrocephalus, seizure disorders such as cerebral palsy and spinal
XX CC bifida, congenital glaucoma, cataract, or sensorineural hearing loss,
XX CC reproductive disorders such as infertility, ovulatory defects and
XX CC endometriosis, autoimmune disorders, ectopic pregnancy and teratogenesis,
XX CC disruption of spermatogenesis, immunological disorders such as AIDS,
XX CC Addison's disease, allergies, bronchitis, atherosclerosis, diabetes
XX CC mellitus, Chron's disease, lupus and irritable bowel syndrome, multiple
XX CC sclerosis, viral, fungal, helminthic, parasitic and protozoal infections,
XX CC and neoplastic disorders including adenocarcinoma, leukaemia, lymphoma,
XX CC melanoma, myeloma, sarcoma, and teratocarcinoma.
XX
XX Sequence 245 AA;

Query Match 100.0%; Score 68; DB 20; Length 245;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VPGAEECERAVTD 13
Db 205 vpgaeceravid 217
|||||

RESULT 4
Y42184 ID Y42184 standard; Protein; 271 AA.
XX AC Y42184;

DT 17-DEC-1999 (first entry)
XX DE Human mFLINT #1 protein sequence.

XX Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
KW reperfusion-associated injury; aplastic anaemia; differentiation;
KW growth; myelodysplastic syndrome; pancytopenic condition;
KW myocardial ischaemia.

XX OS Homo sapiens.

XX PN WO950413-A2.

XX PD 07-OCT-1999.

XX PF 30-MAR-1999; 99WO-US06797.

XX PR 30-MAR-1998; 98US-0079856.
XX PR 20-MAY-1998; 98US-0086074.
XX PR 09-SEP-1998; 98US-0099643.
XX PR 17-DEC-1998; 98US-0112577.
XX PR 18-DEC-1998; 98US-0112703.
XX PR 18-DEC-1998; 98US-0112933.
XX PR 22-DEC-1998; 98US-0113407.
XX PA (ELIL) LILLY & CO ELI.
XX
XX Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
XX PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
XX PI Song HV, Wang J, Wu X, Zuckerman SH;
XX WPI; 1999-591319/50.
XX DR N-PSDB; Z25377.
XX
XX Use of mature FLINT for treating acute liver failure, inflammation,
XX PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
XX PT and proinflammatory activity
XX PS Claim 31; Fig 3; 99pp; English.
XX
XX The present invention describes therapeutic applications of mature FLINT
XX CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
XX CC (mFLINT), which is a member of the tumour necrosis factor receptor
XX CC superfamily, is used for treating acute liver failure, inflammation of
XX CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
XX CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
XX CC injury or disorder such as hypercoagulation (including use with
XX CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
XX CC or disorder, Type I diabetes, cancer, cell damage or damage to an
XX CC innocent bystander tissue that is induced by a chemotherapeutic agent or
XX CC therapeutic irradiation, treating haematopoietic progenitor cells that
XX CC have been exposed to therapeutic radiation or chemotherapy, aplastic
XX CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
XX CC also used for promoting the growth or differentiation of a haematopoietic
XX CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
XX CC resulting from abnormal myocardial ischaemia. The present sequence
XX CC represents human mFLINT.
XX
XX Sequence 271 AA;

Query Match 100.0%; Score 68; DB 20; Length 271;
Best Local Similarity 100.0%; Pred. No. 0.00014;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VPGAEECERAVTD 13
Db 176 vpgaeceravid 188
|||||

RESULT 5
Y42185 ID Y42185 standard; Protein; 273 AA.
XX AC Y42185;

DT 17-DEC-1999 (first entry)
XX DE Human mFLINT #2 protein sequence.

XX Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
KW reperfusion-associated injury; aplastic anaemia; differentiation;
KW growth; myelodysplastic syndrome; pancytopenic condition;
KW myocardial ischaemia.

XX OS Homo sapiens.

Wed Jan 31 13:44:08 2001

PN WO9950413-A2.
 XX
 PD 07-OCT-1999.
 XX
 PF 30-MAR-1999; 99WO-US06797.
 XX
 PR 30-MAR-1998; 98US-0079856.
 PR 20-MAY-1998; 98US-0086074.
 PR 09-SEP-1998; 98US-0099643.
 PR 17-DEC-1998; 98US-0112577.
 PR 18-DEC-1998; 98US-0112703.
 PR 18-DEC-1998; 98US-0112933.
 PR 22-DEC-1998; 98US-0113407.
 XX
 PA (ELIL) LILLY & CO ELI.
 XX
 PI Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
 PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
 PI Song HY, Wang J, Wu X, Zuckerman SH;
 XX
 DR WPI; 1999-591319/50.
 DR N-PSDB; 225378.
 XX
 XX Use of mature FLINT for treating acute liver failure, inflammation,
 PT cancer, and diabetes by prevention of FasL-Fas mediated apoptotic
 PT and proinflammatory activity
 PT
 PS Example 2; Fig 4; 99pp; English.
 XX
 XX The present invention describes therapeutic applications of mature FLINT
 CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
 CC (mFLINT), which is a member of the tumour necrosis factor receptor
 CC superfamily, is used for treating acute liver failure, inflammation of
 CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
 CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
 CC injury or disorder such as hypercoagulation (including use with
 CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
 CC or disorder, Type I diabetes, cancer, cell damage or damage to an
 CC innocent bystander tissue that is induced by a chemotherapeutic agent or
 CC therapeutic irradiation, treating haematopoietic progenitor cells that
 CC have been exposed to therapeutic radiation or chemotherapy, aplastic
 CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
 CC also used for promoting the growth or differentiation of a haematopoietic
 CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
 CC resulting from abnormal myocardial ischaemia. The present sequence
 CC represents human mFLINT.
 XX
 SQ Sequence 273 AA;
 Query Match 100.0%; Score 68; DB 20; Length 273;
 Best Local Similarity 100.0%; Pred. No. 0.00014;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VPAAEECECAVID 13
 Db 178 vpgaacececaavid 190
 |||||
 RESULT 6
 W66102
 ID W66102 standard; Protein; 300 AA.
 XX
 AC W66102;
 XX
 DT 02-DEC-1998 (first entry)
 XX
 DE Amino acid sequence of tumour necrosis related receptor (TR4).
 XX
 KW Human; tumour necrosis related receptor; TR4; agonist; antagonist;
 KW inhibition; chronic; acute; inflammation; arthritis; septicaemia;
 KW autoimmune disease; transplant rejection; stroke; cancer;
 KW Alzheimer's disease.

XX Homo sapiens.
 OS
 XX EP861850-A1.
 PN
 XX 02-SEP-1998.
 PD
 XX 20-JAN-1998; 98EP-0300382.
 PF
 XX 04-FEB-1997; 97US-0794796.
 PR
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 PA
 XX Emery J, Tan KB, Truneh A, Young PR;
 PI
 XX WPI; 1998-508248/44.
 DR
 DR N-PSDB; V07654.
 XX
 XX New DNA encoding tumour necrosis related receptor - used to treat
 PT and prevent e.g. inflammation, arthritis, septicaemia, autoimmune
 PT diseases, transplant rejection, infection, stroke, ischaemia, ARDS,
 PT restenosis, AIDS, bone disorders and cancer
 PT
 XX Claim 1; Fig 1; 2lpp; English.
 PS
 XX This is the amino acid sequence of the human tumour necrosis related
 CC receptor (TR4), used in the method of the invention. The TR4 protein
 CC or its agonist can be used to treat a subject in need of enhanced
 CC TR4 polypeptide activity. The antagonist is used to inhibit TR4
 CC polypeptide activity. The active agents can be used for the
 CC treatment and prevention of diseases such as chronic and acute
 CC inflammation, arthritis, septicaemia, autoimmune diseases, transplant
 CC rejection, stroke, cancer, Alzheimer's disease.
 CC
 XX Sequence 300 AA;
 Query Match 100.0%; Score 68; DB 19; Length 300;
 Best Local Similarity 100.0%; Pred. No. 0.00015;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VPAAEECECAVID 13
 Db 205 vpgaacececaavid 217
 |||||
 RESULT 7
 W63622
 ID W63622 standard; Protein; 300 AA.
 XX
 AC W63622;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Human tumour necrosis factor receptor-6 alpha protein.
 XX
 KW Human tumour necrosis factor receptor-6 alpha; TNFR-6 alpha; TNFR-6 beta;
 KW endothelial cells; keratinocytes; normal prostate; apoptosis;
 KW prostate tumour tissue.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..30
 FT Protein 31..300
 FT /note= "TNFR-6 alpha"
 FT Region 31..282
 FT /note= "Soluble extracellular domain"
 XX
 PN WO9830594-A2.
 XX
 PD 16-JUL-1998.
 XX

PF 13-JAN-1998; 98WO-US00153.
 XX
 PR 14-JAN-1997; 97US-0035496.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ebner R, Feng P, Gentz RL, N1 J, Ruben SM, Yu G;
 XX
 DR WPI; 1998-399142/34.
 DR N-PSDB; V39085.
 XX
 PT Human tumour necrosis factor receptors 6-alpha and 6-beta - used in
 PT the diagnosis of immune system-related disorder(s)
 XX
 PS Claim 20; Fig 1; 91pp; English.
 CC
 CC The present sequence represents the human tumour necrosis factor
 CC receptor-6 alpha (TNFR-6 alpha) protein. The invention also provides
 CC for the TNFR-6 beta protein (W63623). TNFR-6 alpha and TNFR-6 beta are
 CC members of the tumour necrosis factor receptor (TNFR) family. TNFRs
 CC are expressed in endothelial cells, keratinocytes, normal prostate and
 CC prostate tumour tissue. For a number of disorders of these cells,
 CC particularly of the immune system, substantially altered (whether
 CC increased or decreased) levels of TNFR-6 alpha and/or TNFR-6 beta
 CC expression can be detected, therefore the TNFR-6 alpha and TNFR-6 beta
 CC polypeptides, nucleic acids and antibodies are claimed to be useful in
 CC the diagnosis of such disorders. Mutations of the TNFR-6 alpha and
 CC TNFR-6 beta genes can also be detected. The TNFR polypeptides are
 CC also claimed to be useful for identifying ligands which may be useful
 CC in the treatment of apoptosis related disorders.
 XX
 SQ Sequence 300 AA;

Query Match 100.0%; Score 68; DB 19; Length 300;
 Best Local Similarity 100.0%; Pred. No. 0.00015;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VPGAEECERAVID 13
 Db 205 vpgaeceravid 217
 |||||

RESULT 8
 Y03099
 ID Y03099 standard; Protein; 300 AA.
 XX
 AC Y03099;
 XX
 DT 09-DEC-1999 (first entry)
 XX
 DE Human lung TNF-receptor protein.
 XX
 KW Tumour necrosis factor; TNF; TNF receptor; human; lung; gene therapy;
 KW detection; immunoassay; diagnosis; disease; immune system; tumour;
 KW osteogenic system; cardiovascular system; central nervous system; asthma;
 KW peripheral nervous systems; transplant incompatibility; antitumor;
 KW rheumatoid arthritis; antiasthmatic; antiarthritic.
 XX
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 134..1036
 FT /*tag= a
 FT /product= "TNF-receptor"
 XX
 PN DE19809978-A1.
 XX
 PD 16-SEP-1999.
 XX
 PF 09-MAR-1998; 98DE-1009978.
 XX
 PR 09-MAR-1998; 98DE-1009978.

XX (BADI) BASF AG.
 XX
 PI Kroeger B;
 XX
 DR WPI; 1999-519473/44.
 DR N-PSDB; Z09998.
 XX
 PT New soluble member of tumor necrosis factor receptor family, useful for
 PT identification specific modulators and for treating disease e.g. tumors
 XX
 PS Claim 1; Page 8-9; 10pp; German.
 CC
 CC This invention describes a novel tumour necrosis factor (TNF) receptor
 CC (I) isolated from human lung tissue. (I) is used: (i) to raise specific
 CC antibodies (Ab); (ii) to screen for specific (antagonists or ligands
 CC (A), potential therapeutic agents; and (iii) therapeutically (optionally
 CC expressed from a gene therapy vector) in conditions associated with a
 CC deficit of (I). Ab are used: (a) for qualitative or quantitative
 CC detection of (I) in standard immunoassays (for diagnosis of disease, or
 CC susceptibility, or for monitoring); and (b) as therapeutic inhibitors in
 CC cases where (I) is overexpressed. Nucleic acid (II) that encodes (I) is
 CC used: (A) for recombinant production of (I); (B) also its oligonucleotide
 CC fragments, in standard hybridization and/or amplification assays; (C) as
 CC source of antisense molecules or ribozymes; and (D) to produce transgenic
 CC animals (for studying (patho)physiology of (I)). Diseases possibly
 CC associated with under- or over-expression of (I) are those of the immune,
 CC osteogenic, cardiovascular and central or peripheral nervous systems,
 CC tumors, transplant incompatibility, asthma and rheumatoid arthritis. The
 CC products of the invention have antitumor, antiasthmatic and
 CC antiarthritic activity. This sequence represents the TNF-receptor of the
 XX
 SQ Sequence 300 AA;

Query Match 100.0%; Score 68; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 0.00015;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VPGAEECERAVID 13
 Db 205 vpgaeceravid 217
 |||||

RESULT 9
 Y42182
 ID Y42182 standard; Protein; 300 AA.
 XX
 AC Y42182;
 XX
 DT 17-DEC-1999 (first entry)
 XX
 DE Human FLINT #1 protein sequence.
 XX
 KW Human; FLINT; mFLINT; OP63; tumour necrosis factor receptor; FasL;
 KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
 KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
 KW reperfusion-associated injury; aplastic anaemia; differentiation;
 KW growth; myelodysplastic syndrome; pancytopenic condition;
 KW myocardial ischaemia.
 XX
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 134..1036
 FT /*tag= a
 FT /product= "TNF-receptor"
 XX
 PN WO9950413-A2.
 XX
 PD 07-OCT-1999.
 XX
 PF 30-MAR-1999; 99WO-US06797.
 XX
 PR 30-MAR-1998; 98US-0079856.
 PR 20-MAY-1998; 98US-0086074.

```

PR 17-FEB-1998; 98US-0074896.
PR 24-NOV-1997; 97US-0066446.
XX
XX
XX PA (BIOJ ) BIOGEN INC.
XX
XX PI Tschopp J;
XX
XX DR WPI; 1999-347693/29.
XX DR N-PSDB; X76052.
XX
XX PT New tumour necrosis factor family receptor OPG-2
XX
XX PS Claim 1; Page 18; 22pp; English.
XX
XX CC The present sequence represents a mammalian tumour necrosis factor
XX CC receptor, designated OPG-2. OPG-2, is a member of the tumour necrosis
XX CC factor receptor family, and can be used: (i) to raise specific with
XX CC antibodies (Ab), (ii) to treat osteopenic disorders associated with
XX CC excessive osteoclast activity, e.g. primary osteoporosis, Paget's
XX CC disease, hyperglycaemia of malignancy, or osteolytic metastases; (iii)
XX CC for affinity purification of cognate ligands, and (iv) to screen for
XX CC ligands (antagonists or agonists). Ab, or other OPG-2 blocking agents
XX CC such as soluble forms of the protein, are used to prevent, or reduce
XX CC severity of, an immune response, and for treating cancer. They can also
XX CC be used in diagnostic assays. The nucleic acid sequence encoding OPG-2
XX CC can be used as a probe to isolate related sequences from other species.
XX
XX SQ Sequence 300 AA;
XX
XX Query Match 100.0%; Score 68; DB 20; Length 300;
XX Best Local Similarity 100.0%; Pred. No. 0.00015;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX OY 1 VPGAECCECAVID 13
XX | | | | | | | | | |
XX DB 205 vpgaeccecavid 217
XX
XX RESULT 11
XX Y06817
XX ID Y06817 standard; Protein; 300 AA.
XX
XX AC Y06817;
XX
XX DT 24-JUN-1999 (first entry)
XX
XX DE Human Dcr3 polypeptide.
XX
XX KW Dcr3 polypeptide; tumour necrosis factor receptor; TNFR; Fas ligand;
XX KW apoptosis; T cell mediated immune response; allergy; asthma; cancer;
XX KW rheumatoid arthritis; Crohn's disease; guest vs. host disease; human;
XX KW gene therapy.
XX
XX OS Homo sapiens.
XX
XX PN WO9914330-A1.
XX
XX PD 25-MAR-1999.
XX
XX PF 18-SEP-1998; 98WO-US19661.
XX
XX PR 30-JUL-1998; 98US-0094640.
XX PR 18-SEP-1997; 97US-0059288.
XX
XX PA (GETH ) GENENTECH INC.
XX
XX PI Askenazi AJ, Botstein D, Dodge KH, Goddard A, Gurney AL;
XX PI Kim KJ, Lawrence DA, Pitti R, Roy MA, Tumas DB;
XX PI Wood WI;
XX
XX DR WPI; 1999-244032/20.
XX DR N-PSDB; X32744.

```

XX DcR3 polypeptide related to tumor necrosis factor receptor
PT
XX
PS Claim 5; Fig 1: 88pp; English.

This represents a human Dcr3 polypeptide, a homologue of tumour necrosis factor receptor (TNFR) polypeptide. Host cells containing a vector comprising the Dcr3 nucleic acid can be used for the recombinant expression of the protein. Dcr3 binds to Fas ligand, so it (or its chimeras) are useful for modulating apoptosis in mammalian cells, also other Fas-ligand induced activities, particularly to inhibit T cell mediated immune responses, e.g. in treatment of allergy, asthma, rheumatoid arthritis, Crohn's disease, guest vs. host disease etc. Dcr3 may also be used to identify specific binding proteins, potential inhibitors. Antibodies against Dcr3 are used to treat cancer, specifically of the lung and colon, also in diagnosis and for affinity purification of the protein. Detecting mutations in the gene for Dcr3 is also used to diagnose cancer, or predisposition to it. Dcr3 nucleic acid is useful as hybridization probe to detect genomic or related sequences; for chromosome and gene mapping; as source of antisense sequences; for expression of recombinant Dcr3 and to generate transgenic animals (for development and screening of therapeutic agents), also for *in vivo* or *ex vivo* gene therapy.

Sequence 300 AA;

```
Query Match      100.0%; Score 68; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 13; Conservative 0; Mismatches 0; Indels
```

QY	1	VP	GA	EE	CERA	VID	13
Db	205	vp	ga	ee	ce	ra	vid 21

RESULT 12
W97749
ID W97749 standard; Protein; 300 AA.
XX
XX W97749;
XX
XX
XX 21-MAY-1999 (first entry)
XX
XX Human tumour necrosis factor receptor ZTNFR-5.
DE
XX
XX ZTNFR-5; tumour necrosis factor receptor; TNFR; human;
KW cell maturation; bone cell regulation.

AA	
PN	WO9904001-A1.
XX	
PD	28-JAN-1999.
XX	
PF	21-JUL-1998; 98WO-US15072.

PR 21-JUL-1997; 97US-0053203
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Farrah TM;

Novel tumour necrosis factor receptor 2TNFR5 - useful for regulating maturation of TNF-ligand bearing cells

This polypeptide comprises a new, secreted tumour necrosis factor receptor (see W97749), designated zTNFR-5. Novel zTNFR-5 encoding polynucleotides and polypeptides were initially identified by querying an expressed sequence tag (EST) database for sequences homologous to conserved motifs within the TNF receptor family. Based on this search, a contig of 16 ESTs (see X07226) was constructed. zTNFR-5 polypeptides comprise 4 cysteine-rich repeats (see also W97750-55) that are homologous to other TNF receptors, in particular the soluble, secreted TNF receptor osteoprotegerin. zTNFR-5 polypeptide can be prepared by recombinant methods. The polypeptide, especially the extracellular domain, can be used to generate a soluble variant of zTNFR-5. The polypeptides and nucleic acids can be used to screen for ligands, agonists and antagonists of zTNFR-5. The polypeptides can be used in bone cell regulation and to regulate the maturation of TNF ligand-bearing cells such as T- or B-cells, lymphocytes, peripheral blood mononuclear cells, polymorphonuclear leukocytes, fibroblasts or haematopoietic cells.

AA	Sequence	300 AA;
SQ		

```
Query Match      100.0%; Score 68; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. NO. 0.00015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 VPGAEECERAVID 13
Db 205 vpgaeecearavid 217

RESULT 13
W95082
ID W95082 standard: Protein: 300 AA:

DE Orphan receptor (HUMAN NTR-1) polypeptide.

AA HUMAN NTR-1; orphan receptor; osteoprotegerin; OPG; TNFR; human;
KW KW tumour necrosis factor receptor; muscle disorder; bone mass; screening;
KW KW muscle metabolism; binding agent; cognate ligand.

XX Homo sapiens.
OS
XX
PN WO9907738-A2.

XX
PD 18-FEB-1999.

AA 04-AUG-1998; 98WO-US16202.
PF

AA 06-AUG-1997; 97US-0054869.

PA (PROC) PROCTER & GAMBLE CO.

XX
OUT FORM 1 NOVEMBER 1967 (REV) 11

PI Masiakowski PJ, Morris J, Valenzuela DM;

XX WPI; 1999-167365/14.
 DR N-PSDB; X22300.
 XX
 PT Novel orphan human receptor polypeptide and nucleic acid - useful as
 PT diagnostic reagents and for treatment of muscle disorders
 XX
 PS Claim 7; Page 21; 23pp; English.
 XX
 CC This represents a HUMAN NTR-1 polypeptide, a novel orphan receptor. The
 CC protein is related to osteoprotegerin (OPG) and to tumour necrosis factor
 CC receptor (TNFR). Host cells transformed with a vector comprising the
 CC HUMAN NTR-1 nucleic acid are used for the recombinant expression of the
 CC protein. HUMAN NTR-1 proteins and antibodies immuno specific for the
 CC protein are useful for diagnosis and treatment of humans and animals,
 CC especially muscle disorders, as the receptor is involved in regulation of
 CC bone mass and muscle metabolism. HUMAN NTR-1 receptors are also useful
 CC for screening for novel binding agents, and cognate ligands, which may be
 CC used to treat disorders associated with HUMAN NTR-1 imbalance.
 XX
 SQ Sequence 300 AA;
 Query Match 100.0%; Score 68; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 0.00015;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VPGAEECERAVID 13
 DB 205 vpgaeceravid 217
 |||||||||||
 RESULT 14
 Y77458 ID Y77458 standard; Protein; 300 AA.
 XX AC Y77458;
 XX 05-JUN-2000 (first entry)
 DT Human TNF receptor-like protein, HDTEA84.
 DE
 DE
 DE
 XX TNF receptor family; tumour necrosis factor; HDTEA84; HSLJD37R;
 KW Rank-like protein; RANKL; immune disorder; inflammation; allergy;
 KW immunosuppressant; antirheumatic; antirheumatoid; antiinflammatory;
 KW dermatological; antithyroid.
 XX
 XX Homo sapiens.
 OS
 XX WO200001817-A2.
 PN
 XX 13-JAN-2000.
 PD
 XX 06-JUL-1999; 99WO-US12366.
 XX
 XX 06-JUL-1998; 98US-0110938.
 PR 13-JUL-1998; 98US-0114466.
 PR 23-JUL-1998; 98US-0093897.
 PR 12-AUG-1998; 98US-0132968.
 PR 18-AUG-1998; 98US-0136214.
 PR 11-SEP-1998; 98US-0099999.
 XX
 XX (SCHE) SCHERING CORP.
 PA
 XX Bates EEM, Lebecque SJE, Murphy EF, Mattson JD, Gorman DM;
 PI Hedrick JA, Wang L, Zlotnik A, Murgolo NJ, Greene JR, Johnston JA;
 PI Bazan JF, Mahony D, Lees EM;
 XX
 XX WPI; 2000-171015/15.
 DR N-PSDB; 292404.
 DR
 XX New isolated mammalian genes, used to develop products for treating
 PT e.g. immune, inflammatory or allergic abnormalities, cancers or

PT degenerative conditions -
 XX Claim 24; Page 157; 218pp; English.
 PS
 XX
 CC The invention relates to a number of primate and/or rodent proteins, and
 CC the genes which encode them. The invention encompasses human dendritic
 CC cell prostaglandin transporter (DC-PGT); the TNF (tumour necrosis
 CC factor) receptor family-related proteins HDTEA84, HSLJD37R and RANKL;
 CC human CC chemokine HCC5; human deubiquitinating proteins DUB1 and DUB
 CC 12; human MD-1 and human and murine MD-2 proteins, which exhibit the
 CC properties of ligands for proteins comprising a leucine-rich motif (LRR);
 CC human cyclin E2; cDNAs encoding these proteins; and antibodies against
 CC these proteins. The proteins can be used for modulating the physiology or
 CC development of a cell. They can be used to mediate uptake of substrates
 CC (e.g., prostaglandin-like molecules), to modulate or mediate cellular
 CC interactions (e.g., induce or prevent trafficking, proliferation, or
 CC differentiation of cells), or are intracellular proteins which are
 CC important in various cellular processes such as the deubiquitination of
 CC proteins or cell cycle regulation. The products can be used for
 CC treating medical conditions such as immune, inflammatory or allergic
 CC disorders, or abnormal cellular proliferation, for example, cancers or
 CC degenerative conditions. They can be used to modulate immune responses in
 CC disease states e.g., autoimmune disorders, Hashimoto's autoimmune
 CC arthritis, systemic lupus erythematosus, Hashimoto's autoimmune
 CC thyroiditis, as well as acute and chronic inflammatory responses in which
 CC T cell activation, expansion, and/or immunological T cell memory play an
 CC important role. Sequences Y77458-Y77461 and Y77465-Y77468 represent
 CC TNF receptor family-related proteins. Y77458 is the human protein
 CC HDTEA84, Y77459-Y77461 are human HSLJD37R proteins, Y77465 is
 CC murine Rank-like protein RANKL, and Y77466-Y77468 are human RANKL
 CC proteins.
 XX
 SQ Sequence 300 AA;
 Query Match 100.0%; Score 68; DB 21; Length 300;
 Best Local Similarity 100.0%; Pred. No. 0.00015;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VPGAEECERAVID 13
 DB 205 vpgaeceravid 217
 |||||||||||
 RESULT 15
 Y42183 ID Y42183 standard; Protein; 302 AA.
 XX AC Y42183;
 XX 17-DEC-1999 (first entry)
 DT Human FLINT #2 protein sequence.
 DE
 DE
 XX Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; Fasli;
 KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
 KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
 KW reperfusion associated injury; aplastic anaemia; differentiation;
 KW growth; myelodysplastic syndrome; pancytopenic condition;
 XX myocardial ischaemia.
 XX
 OS Homo sapiens.
 XX
 XX WO9950413-A2.
 PN
 XX 07-OCT-1999.
 PD
 XX 30-MAR-1999; 99WO-US06797.
 PF
 XX 30-MAR-1998; 98US-0079856.
 PR 20-MAY-1998; 98US-0086074.
 PR 09-SEP-1998; 98US-0099643.
 PR 17-DEC-1998; 98US-0112577.
 PR


```

PR 18-DEC-1998; 98US-0112703.
PR 18-DEC-1998; 98US-0112933.
PR 22-DEC-1998; 98US-0113407.
XX
XX (ELIL ) LILLY & CO ELI.
XX
XX Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
PI Song HY, Wang J, Wu X, Zuckerman SH;
XX
XX WPI; 1999-591319/50.
DR N-PSDB; 225376.
XX
XX Use of mature FLINT for treating acute liver failure, inflammation,
PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
PT and proinflammatory activity
XX
XX Example 2: Fig 2; 99pp; English.
XX
XX The present invention describes therapeutic applications of mature FLINT
CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
CC (mFLINT), which is a member of the tumour necrosis factor receptor
CC superfamily, is used for treating acute liver failure, inflammation of
CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
CC injury or disorder such as hypercoagulation (including use with
CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
CC or disorder, Type I diabetes, cancer, cell damage or damage to an
CC innocent bystander tissue that is induced by a chemotherapeutic agent or
CC therapeutic irradiation, treating haematopoietic progenitor cells that
CC have been exposed to therapeutic radiation or chemotherapy, aplastic
CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
CC also used for promoting the growth or differentiation of a haematopoietic
CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
CC resulting from abnormal myocardial ischaemia. The present sequence
CC represents human FLINT.
XX
SQ Sequence 302 AA;

```

```

Query Match 100.0%; Score 68; DB 20; Length 302;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPGAEECERAVID 13
Db 207 vpgaeececeravid 219

```

Search completed: January 30, 2001, 16:45:24
Job time: 584 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:50:24 ; Search time 149.64 Seconds
(without alignments)
5.899 Million cell updates/sec

Title: US-09-518-931-2_COPY_205_217
Perfect score: 68
Sequence: 1 VPGAEECERAVID 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_66: *
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	66.2	137	2 A83222	probable ring-clea
2	45	66.2	437	2 A40037	protein farnesyltr
3	41	60.3	431	1 A39586	NADH dehydrogenase
4	40	58.8	243	2 T12485	hypothetical prote
5	40	58.8	375	2 T09374	hypothetical prote
6	40	58.8	596	2 G82874	conserved hypothet
7	40	58.8	716	2 T37710	hypothetical prote
8	39	57.4	152	2 G8298	hypothetical prote
9	39	57.4	340	2 S6311	probable membrane
10	39	57.4	389	2 T30454	hypothetical prote
11	39	57.4	863	2 S06017	neuraxin - rat
12	39	57.4	969	2 T27997	hypothetical prote
13	39	57.4	2364	2 A56577	microtubule-associ
14	39	57.4	2464	1 QRMSP1	microtubule-associ
15	38	55.9	192	2 D71902	hypothetical prote
16	38	55.9	363	2 B82090	acetyltransferase,
17	38	55.9	446	1 A46621	corrinoid/iron-sul
18	38	55.9	558	2 I56545	glypican precursor
19	38	55.9	2124	2 T01526	hypothetical prote
20	37	54.4	117	2 T73315	hypothetical prote
21	37	54.4	215	2 B70622	probable ribosomal
22	37	54.4	250	2 T36428	hypothetical prote
23	37	54.4	319	2 J4390	bone marrow stroma
24	37	54.4	1941	2 T23979	hypothetical prote
25	37	54.4	1943	2 T23986	hypothetical prote
26	36	52.9	85	1 B30552	T-cell activation
27	36	52.9	234	2 C75368	conserved hypothet
28	36	52.9	256	2 E83287	hypothetical prote
29	36	52.9	294	2 A70621	probable argB prot

ALIGNMENTS

RESULT 1

A83222

probable ring-cleaving dioxygenase PA3389 [Imported] - Pseudomonas aeruginosa (strain C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 15-Sep-2000
C:Accession: A83222

R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: A82950
A:Accession: A83222
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-137 <STO>

A:Cross-references: GB:AE004760; GB:AE004091; NID:g9949520; PIDN:AAG06778.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA3389

Query Match 66.2%; Score 45; DB 2; Length 137;
Best Local Similarity 61.5%; Pred No. 0.74; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPGAEECERAVID 13
|| ||||| :|
Db 21 VPNLEECERFYVD 33

RESULT 2

A40037

protein farnesyltransferase (EC 2.5.1.1-) beta chain - rat
C:Species: Rattus norvegicus (Norway rat)

C:Date: 16-Oct-1992 #sequence_revision 16-Oct-1992 #text_change 05-Nov-1999

C:Accession: A40037

R:Chen, W.J.; Andres, D.A.; Goldstein, J.L.; Russell, D.W.; Brown, M.S.

Cell 66, 327-334, 1991

A:Title: cDNA cloning and expression of the peptide-binding beta subunit of rat p21(r

A:Reference number: A40037; MUID:91309145

A:Accession: A40037

A:Molecule type: mRNA

A:Residues: 1-437 <CHE>

A:Cross-references: GB:M69056; NID:g204185; PIDN:AAA41176.1; PID:g204186

C:Comment: This protein attaches farnesyl residues to a cysteine near the carboxyl te

C:Superfamily: DPRI protein

C:Keywords: heterodimer; transferase

Query Match 66.2%; Score 45; DB 2; Length 437;
Best Local Similarity 81.8%; Pred. No. 2.3;

Wed Jan 31 13:44:10 2001

Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VPGAECERAV 11
||| ||||| ||

Db 420 VPGFECEDAV 430

RESULT 3

A39588
NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) I chain 1 - Paracoccus denitrificans
N:Alternate names: NADH dehydrogenase (ubiquinone) 50K chain; NADH dehydrogenase (ubiqui
C:Species: Paracoccus denitrificans
C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 11-Jun-1999
C:Accession: A39588; S23946; D40296
R:Yu, X.; Matsuno-Yagi, A.; Yagi, T.
Biochemistry 30, 6422-6428, 1991
A:Title: The NADH-binding subunit of the energy-transducing NADH-ubiquinone oxidoreducta
A:Reference number: A39588; MUID:91274292
A:Accession: A39588
A:Molecule type: DNA
A:Residues: 1-431 <XUA>
A:Cross-references: GB:M64432; NID:g150597; PIDN:AAA25585.1; PID:g150598
A:Note: parts of this sequence, including the amino end of the mature protein, were dete
R:Yu, X.; Matsuno-Yagi, A.; Yagi, T.
Arch. Biochem. Biophys. 296, 40-48, 1992
A:Title: Structural features of the 66-kDa subunit of the energy-transducing NADH-ubiqui
A:Reference number: S23946; MUID:91363357
A:Accession: S23946
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 416-431 <XU2>
R:Yu, X.; Matsuno-Yagi, A.; Yagi, T.
Biochemistry 30, 8678-8684, 1991
A:Title: Characterization of the 25-kilodalton subunit of the energy-transducing NADH-ub
A:Reference number: A40296; MUID:91363357
A:Accession: D40296
A:Molecule type: DNA
A:Residues: 1-8 <XU3>
A:Cross-references: GB:J05337
C:Genetics:
A:Gene: NQO1
C:Complex: heteromultimer containing about 10 types of chains
C:Function:
A:Description: catalyzes the reduction of ubiquinone to ubiquinol by NADH
A:Pathway: oxidative phosphorylation
C:Superfamily: NADH dehydrogenase (ubiquinone) chain F: NADH dehydrogenase (ubiquinone)
C:Keywords: 4Fe-4S; flavoprotein; FMN; iron-sulfur protein; metalloprotein; NAD; oxidati
F:36-414/Domain: NADH dehydrogenase (ubiquinone) I chain F homology <NUOF>
F:54-63/Region: NAD binding motif
F:167-182/Region: FMN binding motif
F:346,349,352,392/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 60.3%; Score 41; DB 1; Length 431;
Best Local Similarity 38.1%; Pred. No. 12;
Matches 8; Conservative 4; Mismatches 1; Indels 8; Gaps 1;

QY 1 VPG-----AECERAVD 13
||| ||| |||

Db 285 IPGASCPLVTAQCENAIMD 305

RESULT 4

T12485
hypothetical protein DKFZp56402082.1 - human
C:Species: Homo sapiens (man)
C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 28-Jul-2000
C:Accession: T12485
R:Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, June 1999
A:Reference number: 217525
A:Accession: T12485

Query Match 58.8%; Score 40; DB 2; Length 596;
Best Local Similarity 61.5%; Pred. No. 24;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-243 <POU>
A:Cross-references: EMBL:AL080166
A:Experimental source: fetal brain; clone DKFZp56402082
C:Genetics:
A:Note: DKFZp56402082.1
C:Superfamily: Schizosaccharomyces pombe hypothetical protein SPBC3E7.13c

Query Match 58.8%; Score 40; DB 2; Length 243;
Best Local Similarity 61.5%; Pred. No. 10;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 VPGAECERAVD 13
||| ||| |||

Db 173 VPSTEIDRWVID 185

RESULT 5

T09374
hypothetical protein F23K16.210 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 24-Nov-1999
C:Accession: T09374
R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.
submitted to the Protein Sequence Database, June 1999
A:Reference number: Z16652
A:Accession: T09374
A:Molecule type: DNA
A:Residues: 1-375 <BEV>
A:Cross-references: EMBL:AL078620; GSPDB:GN00062; ATSP:F23K16.210
A:Experimental source: cultivar Columbia; BAC clone F23K16
C:Genetics:
A:Gene: ATSP:F23K16.210
A:Map position: 4
C:Superfamily: caffeoyl-CoA 3-O-methyltransferase

Query Match 58.8%; Score 40; DB 2; Length 375;
Best Local Similarity 70.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 PGAECECERAV 11
||| |||||

Db 211 PGTCERCERU 220

RESULT 6

G82874
conserved hypothetical U0570 [imported] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C:Accession: G82874
R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to GenBank, February 2000
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a
A:Reference number: A82870
A:Accession: G82874
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-596 <GLA>
A:Cross-references: GB:AE002155; GB:AF222894; NID:g6899572; PIDN:AAF30984.1; GSPDB:GN
A:Experimental source: serovar 3; biovar 1
C:Genetics:
A:Gene: U0570
A:Genetic code: SGC3
C:Superfamily: conserved hypothetical protein MGL39

Query Match 58.8%; Score 40; DB 2; Length 596;
Best Local Similarity 61.5%; Pred. No. 24;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 VPGAECERAVID 13
:||| |||||
Db 202 IPGFSESERVID 214

RESULT 7
T37710
hypothetical protein SPAC15A10.10 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 04-Mar-2000
C:Accession: T37710
R:Murphy, L.; Harris, D.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1997
A:Reference number: Z21738
A:Accession: T37710
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-716 <MUR>
A:Cross-references: EMBL:D297208; PIDN:CAB10107.1; GSPDB:GNO00066; SPDB:SPAC15A10.10
A:Experimental source: strain 972h; cosmid c15A10
C:Genetics:
A:Gene: SPDB:SPAC15A10.10
A:Map position: 1
A:Introns: 82/2; 273/3; 309/2; 632/2
C:Superfamily: Schizosaccharomyces pombe hypothetical protein SPAC15A10.10

Query Match 58.8%; Score 40; DB 2; Length 716;
Best Local Similarity 53.8%; Pred. No. 29;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 VPGAECERAVID 13
:||| |||||
Db 310 IPGERVCHRMVID 322

RESULT 8
G59298
hypothetical protein AF0391 - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 18-Aug-2000
C:Accession: G59298
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
A:Reference number: A69250; MUID:98049343
A:Accession: G59298
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-152 <KLE>
A:Cross-references: GB:AE000782; NID:g2689400; PIDN:AAB90848.1; PID:g265024
C:Superfamily: Pyrococcus horikoshii hypothetical protein PH1318

Query Match 57.4%; Score 39; DB 2; Length 152;
Best Local Similarity 72.7%; Pred. No. 9.8;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPGAECERAV 11
|| |||||
Db 56 VPDFECERFV 66

RESULT 9
S56311
probable membrane protein YFR055w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein R020
C:Species: Saccharomyces cerevisiae

C:Date: 02-Sep-1995 #sequence_revision 19-Oct-1995 #text_change 20-Jun-2000
C:Accession: S56311; S62267; S63803
R:Murakami, Y.; Naitou, M.; Hagiwara, H.; Shibata, T.; Ozawa, M.; Sasanuma, S.I.; Sas
submitted to the EMBL Data Library, May 1995
A:Description: Analysis of the nucleotide sequence of chromosome VI from Saccharomyces
A:Reference number: S56186
A:Accession: S56311
A:Molecule type: DNA
A:Residues: 1-340 <MUR>
A:Cross-references: EMBL:D50617; NID:g836685; PIDN:BA09295.1; PID:g836811; MIPS:YFR0
R:Murakami, Y.
submitted to the EMBL Data Library, December 1994
A:Reference number: S62230
A:Accession: S62267
A:Molecule type: DNA
A:Residues: 1-340 <MUW>
A:Cross-references: EMBL:D44597; NID:g871938; PIDN:BA08022.1; PID:g871955
R:Eki, T.; Naitou, M.; Hagiwara, H.; Ozawa, M.; Sasanuma, S.I.; Sasanuma, M.; Tsuchiy
Yeast 12, 149-167, 1996
A:Title: Analysis of a 36.2 kb DNA sequence including the right telomere of chromosome
A:Reference number: S63787; MUID:96287652
A:Accession: S63803
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 173-226 <EKI>
A:Cross-references: EMBL:D44597
C:Genetics:
A:Map position: 6R
A:Note: YFR055w
C:Superfamily: O-succinylhomoserine (thiol)-lyase
C:Keywords: phosphoprotein; pyridoxal phosphate; transmembrane protein
F:81-97/Domain: transmembrane #status predicted <TMM>
F:208/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 57.4%; Score 39; DB 2; Length 340;
Best Local Similarity 54.5%; Pred. No. 21;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPGAECERAV 11
:|||||:
Db 241 LPGAECDQAL 251

RESULT 10
T30454
hypothetical protein ORF104 - Lymantria dispar nuclear polyhedrosis virus
C:Species: Lymantria dispar nuclear polyhedrosis virus, LdMNPV
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jul-2000
C:Accession: T30454
R:Kuzio, J.; Pearson, M.N.; Harwood, S.H.; Funk, C.J.; Evans, J.T.; Slavicek, J.M.; R
Virology 253, 17-34, 1999
A:Title: Sequence and analysis of the genome of a baculovirus pathogenic for Lymantri
A:Reference number: Z20836; MUID:99124785
A:Accession: T30454
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-389 <KUZ>
A:Cross-references: EMBL:AF081810; NID:g3822234; PIDN:AAC70290.1; PID:g3822339
C:Superfamily: OpMNPV hypothetical protein 104

Query Match 57.4%; Score 39; DB 2; Length 389;
Best Local Similarity 58.3%; Pred. No. 24;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 PGAECECAVID 13
| ||| |||:
Db 223 PLVEECRAVN 234

RESULT 11
S06017

Wed Jan 31 13:44:10 2001

us-09-518-931-2_copy_205_217.rpr

```

neuraxin - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 07-Sep-1990 #sequence_revision 03-Mar-1995 #text_change 05-Nov-1999
A:Reference number: A56577
A:Accession: S06017
R:Rienitz, A.; Greeningloh, G.; Hermans-Borgmeyer, I.; Kirsch, J.; Littauer, U.Z.; Prior
EMBO J. 8, 2879-2888, 1989
A:Title: Neuraxin, a novel putative structural protein of the rat central nervous system
A:Reference number: S06017; MUID:90059871
A:Accession: S06017
A:Molecule type: mRNA
A:Residues: 1-863 <RIE>
A:Cross-references: EMBL:X16623; NID:g57826; PIDN:CAA34620.1; PID:g57827
A:Note: the authors translated the codon GCA for residue 723 as His
C:Keywords: brain, spinal cord; tandem repeat; tubulin binding
F:258-461/Region: 17-residue repeats
F:628-741/Region: arginine/lysine-rich

Query Match 57.4%; Score 39; DB 2; Length 863;
Best Local Similarity 66.7%; Pred. No. 52;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PGAECCERAVID 13
| | | | | | | |
Db 787 PAEEPSRAVL 798

RESULT 12
T27997
hypothetical protein ZK792.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C:Accession: T27997; T28009
R:Kershaw, J.
Submitted to the EMBL Data Library, December 1995
A:Reference number: Z20452
A:Accession: T27997
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-969 <WIL>
A:Cross-references: EMBL:X68302; PIDN:CAA92636.2; GSPDB:GN000022; CESP:ZK792.1
A:Experimental source: clone ZK792
R:Dobson, R.
Submitted to the EMBL Data Library, December 1995
A:Reference number: Z20454
A:Accession: T28009
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-969 <WIL2>
A:Cross-references: EMBL:X68303; PIDN:CAA92644.2; GSPDB:GN000022; CESP:ZK792.1
A:Experimental source: Clone ZK809
C:Genetics:
A:Gene: CESP:ZK792.1
A:Map position: 4
A:Introns: 4/3; 17/3; 37/1; 79/3; 101/2; 155/2; 191/1; 377/3; 438/3; 486/2; 515/3; 535/1

Query Match 57.4%; Score 39; DB 2; Length 969;
Best Local Similarity 50.0%; Pred. No. 58;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPGAEECECAVI 12
| | | | | | | |
Db 639 VPGGNKQCAPL 650

RESULT 13
A56577
microtubule-associated protein MAP 1B - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 16-Feb-1997
C:Accession: A56577
R:Zauner, W.; Kratz, J.; Staunton, J.; Feick, P.; Wiche, G.

```

```

Eur. J. Cell Biol. 57, 66-74, 1992
A:Title: Identification of two distinct microtubule binding domains on recombinant ra
A:Reference number: A56577; MUID:92347374
A:Accession: A56577
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2364 <ZAU>
A:Cross-references: GB:X60550
A:Experimental source: brain
A:Note: nucleotide sequence not given; conceptual translation not complete
C:Superfamily: microtubule-associated protein MAP1B

Query Match 57.4%; Score 39; DB 2; Length 2364;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PGAECCERAVID 13
| | | | | | | |
Db 2288 PAEEPSRAVL 2299

RESULT 14
QMSP1
microtubule-associated protein MAP1B - mouse
N:Alternate names: microtubule-associated protein MAP1(X); microtubule-associated pro
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 01-Sep-2000
C:Accession: S07549; S44387; A33645
R:Noble, M.; Lewis, S.A.; Cowan, N.J.
J. Cell Biol. 109, 3367-3376, 1989
A:Title: The microtubule binding domain of microtubule-associated protein MAP1B conta
A:Reference number: A33645; MUID:90094539
A:Accession: S07549
A:Molecule type: mRNA
A:Residues: 1-2464 <NOB>
A:Cross-references: EMBL:X51396; NID:g52999; PIDN:CAA35761.1; PID:g53000
R:Sanchez, C.; Padilla, R.; Paciucci, R.; Zabala, J.C.; Avila, J.
Arch. Biochem. Biophys. 310, 428-432, 1994
A:Title: Binding of heat-shock protein 70 (hsp70) to tubulin.
A:Reference number: S44387; MUID:94234720
A:Accession: S44387
A>Status: preliminary
A:Molecule type: protein
A:Residues: 653-663, IC' <SAN>
C:Superfamily: microtubule-associated protein MAP1B
C:Keywords: microtubule binding; phosphoprotein; tandem repeat
F:589-786/Domain: microtubule binding #status experimental <MTB>
F:589-592, 639-642, 649-652, 655-658, 660-663, 668-671, 674-677, 679-682, 683-686, 687-690, 691
R-K-E/D-X)
F:1861-2064/Region: 17-residue repeats
F:91,116,351,888,1124,1153,1168,1208,1662,1877,1918,2003,2030,2054,2083/Binding site:
F:147,969,1336,1562,1563,1702,1708,1990,2057,2063,2419/Binding site: phosphate (Thr
F:1953/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 57.4%; Score 39; DB 1; Length 2464;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PGAECCERAVID 13
| | | | | | | |
Db 2388 PAEEPSRAVL 2399

RESULT 15
D71902
hypothetical protein jhp0670 - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999
C:Accession: D71902
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.

```

; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557
A:Accession: D71902
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-192 <ARN>
A:Cross-references: GB:AE001499; NID:g4155220; PIDN:AAD06254.1; PID:g415523
A:Experimental source: strain J99
C:Genetics:
A:Gene: jhp0670

Query Match 55.9%; Score 38; DB 2; Length 192;
Best Local Similarity 61.5%; Pred No. 19;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 VPGAEECERAVID I3
||| | | |
Db 121 VPGIEGSEKKVID I33

Search completed: January 30, 2001, 16:50:27
Job time: 717 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 17:05:59 ; Search time 80.04 Seconds
(without alignments)
5.245 Million cell updates/sec

Title: US-09-518-931-2_COPY_205_217
Perfect score: 68
Sequence: 1 VPGAEECERAVID 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	45	66.2	437	1	PFTB_RAT
2	41	60.3	431	1	NOOL_PARDE
3	40	58.8	716	1	YDOA_SCHPO
4	39	57.4	340	1	METC_YEAST
5	39	57.4	545	1	TCPA_SCHMA
6	39	57.4	863	1	NRFX_RAT
7	39	57.4	2464	1	MAPB_MOUSE
8	39	57.4	2468	1	MAPB_HUMAN
9	38	55.9	445	1	ACSC_MOOTH
10	38	55.9	558	1	GLYP_RAT
11	37	54.4	215	1	RL25_MYCTU
12	37	54.4	319	1	BST1_RAT
13	37	54.4	1941	1	YRM8_CAEEL
14	36	52.9	294	1	ARGB_MYCTU
15	36	52.9	411	1	EF1G_TRYCR
16	36	52.9	670	1	TBUD_BURPI
17	35	51.5	130	1	YJ84_YEAST
18	35	51.5	261	1	6PGL_STRCO
19	35	51.5	396	1	TRPB_VIBPA
20	35	51.5	490	1	YNC2_YEAST
21	35	51.5	667	1	OSN3_HUMAN
22	35	51.5	889	1	NODV_BRAJA
23	35	51.5	3396	1	POLG_DENIS
24	34	50.0	175	1	GNFK_ECOLI
25	34	50.0	261	1	YJBA_ECOLI
26	34	50.0	311	1	BST1_MOUSE
27	34	50.0	321	1	ASGX_ECOLI
28	34	50.0	388	1	RIR2_BPT4
29	34	50.0	395	1	LIPG_RAT
30	34	50.0	423	1	CLPX_ECOLI
31	34	50.0	433	1	HEM42_SPINOL
32	34	50.0	558	1	GLYP_HUMAN
33	34	50.0	735	1	YKH8_YEAST

34	34	50.0	1085	1	RBP2_BOVIN	P48820 bos taurus
35	34	50.0	1741	1	RPC1_GIALA	P25202 giardia lam
36	34	50.0	2109	1	PKS1_ASPPA	Q12053 aspergillus
37	34	50.0	2541	1	TALI_HUMAN	Q97490 homo sapien
38	34	50.0	2541	1	TALI_MOUSE	P26039 mus musculu
39	34	50.0	3224	1	RBP2_HUMAN	P49792 homo sapien
40	33.5	49.3	409	1	RHAG_HUMAN	Q02094 homo sapien
41	33	48.5	104	1	SCPL_BACSU	P81099 bacillus su
42	33	48.5	109	1	YBAB_ECOLI	P17577 escherichia
43	33	48.5	311	1	SCL_CHICK	P24899 gallus gall
44	33	48.5	321	1	CXA9_HUMAN	Q9UK14 homo sapien
45	33	48.5	321	1	CXA9_MOUSE	O54851 mus musculu

ALIGNMENTS

RESULT 1
PFTB_RAT
ID PFTB_RAT STANDARD; PRT; 437 AA.
AC Q02293;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE PROTEIN FARNESYLTRANSFERASE BETA SUBUNIT (EC 2.5.1.-) (CAAX
DE FARNESYLTRANSFERASE BETA SUBUNIT) (RAS PROTEINS PRENYLTRANSFERASE)
DE (FTASE-BETA).
GN FNTB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=BRAIN;
RX MEDLINE=91309145; PubMed=1855253;
RA Chen W.-J., Andres D.A., Goldstein J.L., Russell D.W., Brown M.S.;
RT "cDNA cloning and expression of the peptide-binding beta subunit of
RT rat p21ras farnesyltransferase, the counterpart of yeast DPRI/RAM1.";
RL Cell 66:327-334(1991).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS).
RX MEDLINE=97218306; PubMed=9065406;
RA Park H.-W., Boduluri S.R., Moomaw J.F., Casey P.J., Beese L.S.;
RT "Crystal structure of protein farnesyltransferase at 2.25-A
RT resolution."
RL Science 275:1800-1804(1997).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (3.4 ANGSTROMS).
RX MEDLINE=98322062; PubMed=9657673;
RA Long S.B., Casey P.J., Beese L.S.;
RT "Cocrystal structure of protein farnesyltransferase complexed with a
RT farnesyl diphosphate substrate."
RL Biochemistry 37:9612-9618(1998).
CC -!- FUNCTION: THE BETA SUBUNIT IS RESPONSIBLE FOR PEPTIDE-BINDING.
CC -!- COFACTOR: BINDS ONE ZINC ION.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT.
CC -!- SIMILARITY: BELONGS TO THE PROTEIN PRENYLTRANSFERASE BETA SUBUNIT
FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

CC EMBL: M69056; AAA41176.1; -
DR PIR: A40037; A40037.
DR PDB: 1FT1; 18-MAR-98.
DR PDB: 1FT2; 18-NOV-98.
DR INTERPRO: IPR001330; -
DR PFAM: PF00432; prenylttrans; 5.

Wed Jan 31 13:44:11 2001

Transferase; Prenyltransferase; Repeat; Zinc; 3D-structure.
KW DOMAIN 121 374 REPEAT-RICH REGION.
FT REPEAT 121 164 APPROXIMATE.
FT REPEAT 172 215 APPROXIMATE.
FT REPEAT 220 263 APPROXIMATE.
FT REPEAT 268 312 APPROXIMATE.
FT REPEAT 330 374 APPROXIMATE.
FT REPEAT 397 297 ZINC.
FT METAL 299 299 ZINC.
FT METAL 362 362 ZINC.
SQ SEQUENCE 437 AA; 48673 MW; 41A9D6D79CD319A8 CRC64;

Query Match 66.2%; Score 45; DB 1; Length 437;
Best Local Similarity 81.8%; Pred. No. 1;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VPGAEECERAV 11
||| ||||| ||
DB 420 VPGFEECEDAV 430

RESULT 2
ID NOO1_PARDE STANDARD; PRT; 431 AA.
AC P29913;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 1 (EC 1.6.5.3) (NADH
DE DEHYDROGENASE 1, CHAIN 1) (NDH-1, CHAIN 1).
GN NOO1.
OS Paracoccus denitrificans.
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Paracoccus.
RN [1]
RC SEQUENCE FROM N.A., AND SEQUENCE OF 1-37.
RX STRAIN-ATCC 13548;
RA MEDLINE-91274292; PubMed-1905152;
RA Xu X., Matsuno-Yagi A., Yagi T.;
RT "The NADH-binding subunit of the energy-transducing NADH-ubiquinone
RT oxidoreductase of Paracoccus denitrificans: gene cloning and deduced
RT primary structure.";
RL Biochemistry 30:6422-6428(1991).
RN [2]
RC SEQUENCE OF 1-8 FROM N.A., AND SEQUENCE OF 1-37 AND 194-203.
RX STRAIN-ATCC 13543;
RA MEDLINE-91363357; PubMed-1909571;
RA Xu X., Matsuno-Yagi A., Yagi T.;
RT "Characterization of the 25-kilodalton subunit of the
RT energy-transducing NADH-ubiquinone oxidoreductase of Paracoccus
RT denitrificans: sequence similarity to the 24-kilodalton subunit of
RT the flavoprotein fraction of mammalian complex I.";
RL Biochemistry 30:8678-8684(1991).
CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
CC -1- COFACTOR: FMN AND ONE 4FP-4S CLUSTER (BY SIMILARITY).
CC -1- SUBUNIT: COMPOSED OF 14 DIFFERENT SUBUNITS.
CC -1- SIMILARITY: BELONGS TO THE COMPLEX I 51 DA SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M64432; AAA25585.1; -;
DR PIR: A39588; A39588.
DR PIR: D40296; D40296.
DR INTERPRO: IPR001949; -;
DR PFAM: PF01512; Complex1_51K; 1.
DR PROSITE: PS00644; COMPLEX1_51K_1; 1.

DR PROSITE: PS00645; COMPLEX1_51K_2; 1.
KW Oxidoreductase; NAD; Ubiquinone; Flavoprotein; FMN; Iron-sulfur;
KW 4Fe-4S.
FT NP_BIND 54 63 NAD (BY SIMILARITY).
FT NP_BIND 167 214 FMN (BY SIMILARITY).
FT METAL 346 346 IRON-SULFUR (4FE-4S) (POTENTIAL).
FT METAL 349 349 IRON-SULFUR (4FE-4S) (POTENTIAL).
FT METAL 352 352 IRON-SULFUR (4FE-4S) (POTENTIAL).
FT METAL 392 392 IRON-SULFUR (4FE-4S) (POTENTIAL).
SQ SEQUENCE 431 AA; 47191 MW; AA25C5A0A7570048 CRC64;

Query Match 60.3%; Score 41; DB 1; Length 431;
Best Local Similarity 38.1%; Pred. No. 5.3;
Matches 8; Conservative 4; Mismatches 1; Indels 8; Gaps 1;

QY 1 VPG-----AEECERAVID 13
:|| :||| :||
DB 285 IPGGASCPVLTAECENAIMD 305

RESULT 3
ID YDOA_SCHPO STANDARD; PRT; 716 AA.
AC O13730;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DT HYPOTHETICAL 83.6 KDA PROTEIN C15A10.10 IN CHROMOSOME 1.
DE SPAC15A10.10.
GN Schizosaccharomyces pombe (Fission yeast).
OS Schizosaccharomycetes; Ascomycota; Schizosaccharomycetales;
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetaceae; Schizosaccharomycetes.
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN-972;
RA Murphy L., Harris D., Wood V., Barrell B.G., Rajandream M.A.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z97208; CAB10107.1; -;
DR KW Hypothetical protein.
SQ SEQUENCE 716 AA; 83605 MW; BD2F50D06F56400D CRC64;

Query Match 58.8%; Score 40; DB 1; Length 716;
Best Local Similarity 53.8%; Pred. No. 13;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 VPGAEECERAVID 13
:|| :||| :||
DB 310 IPGERVCHRWVID 322

RESULT 4
ID MEVC_YEAST STANDARD; PRT; 340 AA.
AC P43623;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE PUTATIVE CYSTATHIONINE BETA-LYASE (EC 4.4.1.8) (CEL) (BETA-
DE CYSTATHIONASE) (CYSTEINE LYASE).
GN YFR055W
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;

Saccharomycetaceae; Saccharomyces.
[1]
SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA MEDLINE=95400292; PubMed=7670463;
RX Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.;
RT "Analysis of the nucleotide sequence of chromosome VI from
Saccharomyces cerevisiae".
RL Nat. Genet. 10:261-268(1995).
[2]
SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=96287652; PubMed=8686379;
RA Eki T., Naitou M., Hagiwara H., Ozawa M., Sasanuma S.-I.,
RA Sasanuma M., Tsuchiya Y., Shibata T., Hanaoka F., Murakami Y.;
RT "Analysis of a 36.2 kb DNA sequence including the right telomere of
chromosome VI from Saccharomyces cerevisiae".
RL Yeast 12:149-167(1996).
CC -!- CATALYTIC ACTIVITY: CYSTATHIONINE + H(2)O -> L-HOMOCYSTEINE + NH(3)
CC + PYRUVATE.
CC -!- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -!- PATHWAY: THIRD STEP IN METHIONINE BIOSYNTHESIS.
CC -!- SIMILARITY: BELONGS TO THE TRANS-SULFURATION ENZYMES FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; D50617; BAA09295.1; -;
DR EMBL; D44597; BAA08022.1; -;
DR HSSP; P06721; ICL1.
DR SGD; S0001952; YFR055W.
DR INTERPRO; IPR000277; -;
DR PFAM; PF01053; Cys_Met_Meta_PP; 1.
DR PROSITE; PS00868; CYS_MET_META_PP; 1.
KW Hypothetical protein; Methionine biosynthesis; Lyase;
KW Pyridoxal phosphate.
FT BINDING 208 208 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 340 AA; 36971 MW; 75FF4C7BDA038B8F CRC64;

Query Match 57.4%; Score 39; DB 1; Length 340;
Best Local Similarity 54.5%; Pred. NO. 9.5;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 VPGAEECEERAV 11
:|||||:|:
Db 241 LPGAEDCQAL 251

RESULT 5
TCPA_SCHMA STANDARD; PRT; 545 AA.
AC Q94757;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE T-COMPLEX PROTEIN 1, ALPHA SUBUNIT (TCP-1-ALPHA) (CCT-ALPHA).
GN TCP-1A.
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Neodermata; Trematoda; Digenea;
OC Strigeida; Schistosomatidae; Schistosomatidae; Schistosoma.
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=EGYPTIAN ISOLATE;
RA Campos E.G., Hamdan F.F., Ribeiro P., Prichard R.K.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.

-!- FUNCTION: MOLECULAR CHAPERONE; ASSIST THE FOLDING OF PROTEINS UPON
ATP HYDROLYSIS. KNOWN TO PLAY A ROLE, IN VITRO, IN THE FOLDING OF
ACTIN AND TUBULIN (BY SIMILARITY).
-!- SUBUNIT: HETERO-OLIGOMERIC COMPLEX OF ABOUT 850 TO 900 KDA THAT
FORMS TWO STACKED RINGS, 12 TO 16 NM IN DIAMETER.
-!- SUBCELLULAR LOCATION: CYTOPLASMIC.
-!- SIMILARITY: BELONGS TO THE TCP-1 CHAPERONIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; U55769; AAA99815.1; -;
DR HSSP; P48424; IASX.
DR INTERPRO; IPR002194; -;
DR PFAM; PF00118; cpn60_TCP1; 1.
DR PRINTS; PR00304; TCOMPLEXTCP1.
DR PROSITE; PS00750; TCP1_1; 1.
DR PROSITE; PS00751; TCP1_2; 1.
DR PROSITE; PS00995; TCP1_3; 1.
KW Chaperone; ATP-binding; Multigene family.
SQ SEQUENCE 545 AA; 59439 MW; 992CCA919389E42C CRC64;

Query Match 57.4%; Score 39; DB 1; Length 545;
Best Local Similarity 72.7%; Pred. NO. 15;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPGAEECEERAV 11
:|||||:|:
Db 406 VPGAGACETAV 416

RESULT 6
NERX_RAT STANDARD; PRT; 863 AA.
AC P15205;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE NEURAXIN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90059871; PubMed=2555150;
RA Rientz A., Grenningloh G., Hermans-Borgmeyer I., Kirsch J.,
RA Littauer U.Z., Prior P., Gundelfinger E.D., Schmitt B., Betz H.;
RT "Neuraxin, a novel putative structural protein of the rat central
nervous system that is immunologically related to microtubule-
associated protein 5".
RL EMBO J. 8:2879-2888(1989).
CC -!- FUNCTION: PUTATIVE STRUCTURAL PROTEIN OF THE CENTRAL NERVOUS
SYSTEM. SPECIFICALLY BIND TUBULIN.
CC -!- SIMILARITY: TO MAP1B.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X16623; CAA34620.1; -;
DR PIR; S06017; S06017.
DR INTERPRO; IPR000102; -;

Wed Jan 31 13:44:11 2001

entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC CC PFAM: PF00414; MAP1B_neuraxin; 10.
CC CC PROSITE; PS00230; MAP1B_NEURAXIN; 8.
CC CC NEURONE; Microtubules; Repeat
DR DR DOMAIN 258 461 12 x 17 AA TANDEM REPEATS.
KW REPEAT 258 274
FT REPEAT 275 291
FT REPEAT 292 308
FT REPEAT 309 325
FT REPEAT 326 342
FT REPEAT 343 359
FT REPEAT 360 376
FT REPEAT 377 393
FT REPEAT 394 410
FT REPEAT 411 427
FT REPEAT 428 444
FT REPEAT 445 461
SQ SEQUENCE 863 AA; 94369 MW; D3AD0DAC55235E5E CRC64;

Query Match 57.4%; Score 39; DB 1; Length 863;
Best Local Similarity 66.7%; Pred. No. 24;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PGAECECERAVID 13
DB 787 PAEEPSRAVID 798

RESULT 7
MAPB_MOUSE STANDARD; PRT; 2464 AA.
AC P14873;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE MICROTUBULE-ASSOCIATED PROTEIN 1B (MAP1.2) [CONTAINS: MAP1
DE LIGHT CHAIN LC1].
DE MAP1B OR MTAP5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-SWISS WEBSTER; TISSUE-BRAIN;
RX MEDLINE=90094539; PubMed=2480963;
RA Noble M., Lewis S.A., Cowan N.J.;
RT "The microtubule binding domain of microtubule-associated protein
RT MAP1B contains a repeated sequence motif unrelated to that of MAP2
RT and tau."
RL J. Cell Biol. 109:3367-3376(1989).
CC -!- FUNCTION: THE FUNCTION OF BRAIN MAPS IS ESSENTIALLY UNKNOWN.
CC PHOSPHORYLATED MAP1B MAY PLAY A ROLE IN THE CYTOSKELETAL CHANGES
CC THAT ACCOMPANY NEURITE EXTENSION. POSSIBLY MAP1B BINDS TO AT LEAST
CC TWO TUBULIN SUBUNITS IN THE POLYMER, AND THIS BRIDGING OF SUBUNITS
CC MIGHT BE INVOLVED IN NUCLEATING MICROTUBULE POLYMERIZATION AND IN
CC STABILIZING MICROTUBULES.
CC -!- SUBUNIT: 3 DIFFERENT LIGHT CHAINS, LC1, LC2 AND LC3, CAN ASSOCIATE
CC WITH MAP1A AND MAP1B PROTEINS.
CC -!- DOMAIN: HAS A HIGHLY BASIC REGION WITH MANY COPIES OF THE SEQUENCE
CC KKEE AND KKEI/V, REPEATED BUT NOT AT FIXED INTERVALS. THIS LATTER
CC REGION IS RESPONSIBLE FOR THE BINDING OF MAP1B TO MICROTUBULES
CC BOTH IN VITRO AND IN VIVO.
CC -!- PTM: LC1 IS COEXPRESSED WITH MAP1B. IT IS A POLYPEPTIDE GENERATED
CC FROM MAP1B BY PROTEOLYTIC PROCESSING. IT IS FREE TO ASSOCIATE WITH
CC BOTH MAP1A AND MAP1B. IT INTERACTS WITH THE AMINO-TERMINAL REGION
CC OF MAP1B.
CC -!- SIMILARITY: TO NEURAXIN.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC CC EMBL; X51396; CAA35761.1;
CC CC PIR; S07549; QRMSPL.
CC CC MGI; MGI:97179; MTAP5.
CC CC INTERPRO; IPR000102;
CC CC PFAM; PF00414; MAP1B_Neuraxin; 10.
CC CC PROSITE; PS00230; MAP1B_NEURAXIN; 7.
CC CC Microtubules; Repeat; Phosphorylation.
KW CHAIN ? 2464 MAP1 LIGHT CHAIN LC1.
FT DOMAIN 589 787 KKEE AND KKEI/V REPEATS.
FT DOMAIN 1865 2068 12 x 17 AA TANDEM REPEATS.
FT REPEAT 1865 1881
FT REPEAT 1882 1898
FT REPEAT 1899 1915
FT REPEAT 1916 1932
FT REPEAT 1933 1949
FT REPEAT 1950 1966
FT REPEAT 1967 1983
FT REPEAT 1984 2000
FT REPEAT 2001 2017
FT REPEAT 2018 2034
FT REPEAT 2035 2051
FT REPEAT 2052 2068
SQ SEQUENCE 2464 AA; 270408 MW; FB03DD99CFDBDA87 CRC64;

Query Match 57.4%; Score 39; DB 1; Length 2464;
Best Local Similarity 66.7%; Pred. No. 66;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PGAECECERAVID 13
DB 2388 PAEEPSRAVID 2399

RESULT 8
MAPB_HUMAN STANDARD; PRT; 2468 AA.
AC P46821;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE MICROTUBULE-ASSOCIATED PROTEIN 1B [CONTAINS: MAP1 LIGHT CHAIN LC1].
DE MAP1B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=95104835; PubMed=7806212;
RX Lien L.L., Feener C., Fischbach N., Kunkel L.M.;
RA "Cloning of human microtubule-associated protein 1B and the
RA identification of a related gene on chromosome 15."
RL Genomics 22:273-280(1994).
CC -!- FUNCTION: THE FUNCTION OF BRAIN MAPS IS ESSENTIALLY UNKNOWN.
CC PHOSPHORYLATED MAP1B MAY PLAY A ROLE IN THE CYTOSKELETAL CHANGES
CC THAT ACCOMPANY NEURITE EXTENSION. POSSIBLY MAP1B BINDS TO AT LEAST
CC TWO TUBULIN SUBUNITS IN THE POLYMER, AND THIS BRIDGING OF SUBUNITS
CC MIGHT BE INVOLVED IN NUCLEATING MICROTUBULE POLYMERIZATION AND IN
CC STABILIZING MICROTUBULES.
CC -!- SUBUNIT: 3 DIFFERENT LIGHT CHAINS, LC1, LC2 AND LC3, CAN ASSOCIATE
CC WITH MAP1A AND MAP1B PROTEINS.
CC -!- DOMAIN: HAS A HIGHLY BASIC REGION WITH MANY COPIES OF THE SEQUENCE
CC KKEE AND KKEI/V, REPEATED BUT NOT AT FIXED INTERVALS. THIS LATTER
CC REGION IS RESPONSIBLE FOR THE BINDING OF MAP1B TO MICROTUBULES
CC BOTH IN VITRO AND IN VIVO.
CC -!- PTM: LC1 IS COEXPRESSED WITH MAP1B. IT IS A POLYPEPTIDE GENERATED
CC FROM MAP1B BY PROTEOLYTIC PROCESSING. IT IS FREE TO ASSOCIATE WITH
CC BOTH MAP1A AND MAP1B.
CC -!- SIMILARITY: TO NEURAXIN.

```

```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L06237; AAA18904.1;
DR MIM; 157129;
DR INTERPRO; IPR000102;
DR PFM; PF00414; MAP1B_NEURAXIN; 10.
DR PROSITE; PS00230; MAP1B_NEURAXIN; 6.
KW Microtubules; Repeat; Phosphorylation.
FT CHAIN ? 2468
FT DOMAIN 589 790
FT 1869 2074
FT REPEAT 1869 1885
FT REPEAT 1886 1902
FT REPEAT 1903 1919
FT REPEAT 1920 1936
FT REPEAT 1937 1953
FT REPEAT 1954 1970
FT REPEAT 1971 1987
FT REPEAT 1988 2004
FT REPEAT 2005 2021
FT REPEAT 2022 2038
FT REPEAT 2039 2055
FT REPEAT 2056 2072
FT SEQUENCE 2468 AA; 270618 MW; 540839CBDF09D461 CRC64;
CC -----
Query Match 57.4%; Score 39; DB 1; Length 2468;
Best Local Similarity 66.7%; Pred. NO. 66;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 2 PGAECECAVID 13
DB 2392 PAEEPSRAVL 2403
CC -----
RESULT 9
ACSC_MOOPH
ID ACSC_MOOTH STANDARD; PRT; 445 AA.
AC Q07340;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE CORINOID/IRON-SULFUR PROTEIN, LARGE SUBUNIT (C/FE-SP).
GN ACSC.
OS Moorella thermoacetica (Clostridium thermoaceticum).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Moorella group;
OC Moorella.
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE-93194857; PubMed-8449924;
RA Lu W.-P., Schiav I., Cunningham J.R., Ragsdale S.W.;
RT "Sequence and expression of the gene encoding the
RT corrinoid/iron-sulfur protein from Clostridium thermoaceticum and
RT reconstitution of the recombinant protein to full activity.";
RL J. Biol. Chem. 268:5605-5614(1993).
RN [2]
RP SEQUENCE OF 1-7 FROM N.A., AND SEQUENCE OF 1-18.
RC STRAIN-DSM 521;
RX MEDLINE-89098907; PubMed-2911576;
RA Roberts D.L., James-Hagstrom J.E., Garvin D.K., Gorst C.M.,
RA Runquist J.A., Baur J.R., Haase F.C., Ragsdale S.W.;
RT "Cloning and expression of the gene cluster encoding key proteins
RT involved in acetyl-CoA synthesis in Clostridium thermoaceticum: CO
RT dehydrogenase, the corrinoid/Fe-S protein, and methyltransferase.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:32-36(1989).
```

```
CC -----
CC !- FUNCTION: ACTS AS A METHYL GROUP CARRIER IN THE ANAEROBIC ACETYL-
CC COA PATHWAY (WOOD/LJUNGDAHL PATHWAY) OF CARBON MONOXIDE AND CARBON
CC DIOXIDE FIXATION.
CC !- SUBUNIT: HETERODIMER OF A LARGE AND SMALL CHAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L07099; AAA23254.1;
DR PIR; A46621; A46621.
DR HSP; P01088; 1BFA.
KW Cobalt; Iron-sulfur; 4Fe-4S; Carbon dioxido fixation.
FT INIT_MET 0
FT METAL 16 16 IRON-SULFUR (4FE-4S) (PROBABLE).
FT METAL 19 19 IRON-SULFUR (4FE-4S) (PROBABLE).
FT METAL 24 24 IRON-SULFUR (4FE-4S) (PROBABLE).
FT SEQUENCE 445 AA; 48021 MW; 844183C5FFC8D7AA CRC64;
CC -----
Query Match 55.9%; Score 38; DB 1; Length 445;
Best Local Similarity 58.3%; Pred. NO. 19;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 2 PGAECECAVID 13
DB 232 PGARETSRAID 243
CC -----
RESULT 10
GLYP_RAT
ID GLYP_RAT STANDARD; PRT; 558 AA.
AC P35053;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE GLYPICAN-1 PRECURSOR (HSPG M12).
GN GPC1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 24-55 AND 424-445.
RC TISSUE-BRAIN;
RX MEDLINE-93038690; PubMed-1417860;
RA Karthikeyan L., Maurel P., Rauch U., Margolis R.K., Margolis R.U.;
RT "Cloning of a major heparan sulfate proteoglycan from brain and
RT identification as the rat form of glypican.";
RL Biochem. Biophys. Res. Commun. 188:395-401(1992).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 83-112; 196-207 AND 422-443.
RC STRAIN-NEW ENGLAND DECONESS HOSPITAL;
RX MEDLINE-94267529; PubMed-8207484;
RA Litwack E.D., Stipp C.S., Kumbasar A., Lander A.D.;
RT "Neuronal expression of glypican, a cell-surface
RT glycosylphosphatidylinositol-anchored heparan sulfate proteoglycan,
RT in the adult rat nervous system.";
RL J. Neurosci. 14:3713-3724(1994).
CC !- FUNCTION: CELL SURFACE PROTEOGLYCAN THAT BEARS HEPARAN SULFATE.
CC MAY PLAY AN IMPORTANT ROLE IN THE TROPIC AND INJURY RESPONSES OF
CC NEURONS.
CC !- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.
CC !- TISSUE SPECIFICITY: NERVOUS SYSTEM.
CC !- PTM: THIS CELL-ASSOCIATED GLYPICAN IS FURTHER PROCESSED TO GIVE
CC RISE TO A MEDIUM-RELEASED GLYPICAN.
CC !- SIMILARITY: BELONGS TO THE GLYPICAN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/announcement/](http://www.isb-sib.ch/announcement) or send an email to license@isb-sib.ch).

CC	
EMBL; L02896;	AAA86439.1; -.
EMBL; L34067;	AAA41251.1; -.
PIR; JCI281;	JCI281.
INTERPRO;	IPro01863; -.
Pfam; PF01153;	Glypican; 1.
PROSITE; PS01207;	GLYPICAN; 1.
Proteoglycan;	Heparan sulfate;
KW	Glycoprotein; Signal; GPI-anchor;
kw	Extracellular matrix.

FT	SIGNAL	1	23	GLYPICAN-1. REMOVED IN MATURE FORM (POTENTIAL).
FT	CHAIN	24	530	
FT	PROPEP	531	558	
FT				

LIPID	530	530	GPI-ANCHOR (POTENTIAL).
FT			
CARBOHYD	79	79	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT			
CARBOHYD	116	116	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT			
CARBOHYD	55	55	O-LINKED (GLYCOSAMINOGLYCAN (POTENTIAL).
FT			
CARBOHYD			O-LINKED (GLYCOSAMINOGLYCAN (POTENTIAL).
FT			

FT	CARBOHYD	486	O-LINKED (GLYCOSAMINOGLICAN) (POTENTIAL)
FT	CARBOHYD	488	O-LINKED (GLYCOSAMINOGLICAN) (POTENTIAL)
FT	CARBOHYD	488	O-LINKED (GLYCOSAMINOGLICAN) (POTENTIAL)

FT	CARBOHYD	490	O-LINKED (GLYCOSAMINOGLICAN) (POLYMERIZING).
FT	CONFLICT	21	T -> A (IN REF. 2).
FT	CONFLICT	312	Y -> N (IN REF. 2).

FT	CONFLICT	362	A -> G (IN REF. 2).
FT	CONFLICT	362	A -> G (IN REF. 2).
FT	CONFLICT	515	I -> T (IN REF. 2).
FT	CONFLICT	437	I -> G (IN REF. 2; AA SEQUENCE).

FT	CONFLICT	443	E -> D (IN REF. 2; AA SEQUENCE).
SO	SEQUENCE	558 AA: 61734 MW: E2878A854B9A1D7F CRC64;	

COMPTON

Query Match	Score	DB 1;	Length
55.98;	Score 38;	DB 1;	Length 558;

Best Local Similarity	66.7%	Pred. NO. 23;	
Matches	8;	Conservative	1;
Mismatches			3;
Indels			0;
Gaps			0

Qy 1 VPGAEECERAVI 12

Db 240 VPLAECRAVM 251

RESULT	STANDARD	PR.
RL25_MYCTU	RL25_MYCTU	215 AA
ID	ID	

ID	KU23-MIC10	STANDARD)	TIME)
AC	P96385;		
DT	01-NOV-1997	(Rel	35. Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE PROBABLE 50S RIBOSOMAL PROTEIN L25.
GN RPLV OR RV1015C OR MTCV10G2.34.

OS Mycobacterium tuberculosis.

OC Bacteria: Firmicutes: Actinobacteria; Actinobacteridae;

Actinomycetales; Corynebacterineae; Mycobacterium.
[1]

RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;

RA Cole S.T., Brosch R., parkhill J., Garnier T., Churcher C., Harris D.,
MEDLINE=98295987; PubMed=9634230;
RX

RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA

RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,

RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squires R., Squires S., Sulston J.E.,

RA Taylor K., Whitehead S., Barrell B.G. ;
RT "Deciphering the biology of *Mycobacterium tuberculosis* from the

RT complete genome sequence.^a ;
RL Nature 393:537-544(1998).

CC -|- FUNCTION: BINDS TO THE 5S RNA (BY SIMILARITY).
CC -|- SIMILARITY: BELONGS TO THE L25P FAMILY OF RIBOSOMAL PROTEINS.

This swiss-prot entry is copyright. It is produced through a collaboration

```
FT CARBOHYD 193      N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 319 AA; 35131 MW; 46831685DE2B2472 CRC64;

Query Match      54.4%; Score 37; DB 1; Length 319;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 PGAECECAVID 13
DB 159 PTAEDCENNAVD 170

RESULT 13
YRM8_CAEEL
ID YRM8_CAEEL STANDARD; PRT; 1941 AA.
AC Q09417;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL 216.3 KDA PROTEIN R06F6.8 IN CHROMOSOME II.
GN R06F6.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
[1]
RN SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Chui C.;
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; 246794; CA86778.1; -.
DR WORMPEP; R06F6.8; CE01624.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 517 537 POTENTIAL.
FT TRANSMEM 678 698 POTENTIAL.
FT TRANSMEM 789 809 POTENTIAL.
FT TRANSMEM 812 832 POTENTIAL.
FT TRANSMEM 966 986 POTENTIAL.
FT TRANSMEM 1189 1209 POTENTIAL.
FT TRANSMEM 1227 1247 POTENTIAL.
FT TRANSMEM 1555 1575 POTENTIAL.
FT TRANSMEM 1765 1785 POTENTIAL.
FT TRANSMEM 1831 1851 POTENTIAL.
FT DOMAIN 57 61 POLY-GLN.
FT DOMAIN 247 251 POLY-HIS.
SQ SEQUENCE 1941 AA; 216274 MW; 7978F5301D0A4BD CRC64;

Query Match      54.4%; Score 37; DB 1; Length 1941;
Best Local Similarity 54.5%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 PGAECECAVID 12
DB 873 PAMEHCDRAVL 883

RESULT 14
ARGB_MYCTU
ID ARGB_MYCTU STANDARD; PRT; 294 AA.
AC P94989;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
```

```
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ACETYLGLUTAMATE KINASE (EC 2.7.2.8) (NAG KINASE) (AGK) (N-ACETYL-L-
DE GLUTAMATE 5-PHOSPHOTRANSFERASE).
GN ARG OR RV1654 OR MTCY006H11.19.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-H37RV;
RC MEDLINE-98295987; PubMed-9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E.,
RA Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
CC -!- CATALYTIC ACTIVITY: ATP + N-ACETYL-L-GLUTAMATE = ADP +
CC N-ACETYL-L-GLUTAMATE 5-PHOSPHATE.
CC -!- PATHWAY: SECOND STEP IN ARGININE BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE ACETYLGLUTAMATE KINASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; 285982; CAB06648.1; -.
DR TUBERCULIST; RV1654; -.
DR INTERPRO; IPR001048; -.
DR PFAM; PF00696; akinase; 1.
KW Arginine biosynthesis; Transferase; Kinase.
SQ SEQUENCE 294 AA; 30937 MW; 5EFA44E9E3E8CFC7 CRC64;

Query Match      52.9%; Score 36; DB 1; Length 294;
Best Local Similarity 66.7%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 VPGAEECECAVI 12
DB 249 VPKVEACLRAVI 260

RESULT 15
EFIG_TRYCR
ID EFIG_TRYCR STANDARD; PRT; 411 AA.
AC P34715.
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE ELONGATION FACTOR 1-GAMMA (EF-1-GAMMA).
OS Trypanosoma cruzi.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=Y;
RC MEDLINE-93376525; PubMed-8367313;
RA Billaut-Mulot O., Pommier V., Schoenck R., Plumas-Marty B.,
RA Taibi A., Loyens M., Capron A., Ouassii M.A.;
RT "Nucleotide sequence of a Trypanosoma cruzi cDNA encoding a protein
RT homologous to mammalian E1 gamma."
RL Nucleic Acids Res. 21:3901-3901(1993).
CC -!- FUNCTION: PROBABLY PLAYS A ROLE IN ANCHORING THE COMPLEX TO OTHER
```

wed Jan 31 13:44:11 2001

CC CELLULAR COMPONENTS.
CC -I- SUBUNIT: EF-1 IS COMPOSED OF FOUR SUBUNITS: ALPHA, BETA,
CC DELTA, AND GAMMA.
CC -I- SIMILARITY: TO OTHER SPECIES' GAMMA SUBUNIT AS WELL AS TO THE
CC N-TERMINAL OF SOME EUKARYOTIC AMINO ACID-TRNA SYNTHETASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; L17307; AAA02936.1; -.
DR PIR; S41648; S41648.
DR INTERPRO; IPR000521; -.
DR INTERPRO; IPR001662; -.
DR PFAM; PF00647; EFIG_domain; 1.
DR PFAM; PF00043; GST; 1.
DR PROSITE; PS50040; EFIG; 1.
KW Elongation factor; Protein biosynthesis.
SQ SEQUENCE 411 AA; 46567 MW; F61463694B73CD80 CRC64;

Query Match 52.9%; Score 36; DB 1; Length 411;
Best Local Similarity 77.8%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Caps 0;

QY 2 PGAECCERA 10
Db 228 PGREAERA 236

Search completed: January 30, 2001, 17:06:01
Job time: 1222 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: January 30, 2001, 16:54:46 ; Search time 254.27 seconds
(without alignments)
5.992 Million cell updates/sec

Title: US-09-518-931-2_COPY_205_217

Perfect score: 68

Sequence: 1 VPGAECERAVID 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_15:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	100.0	300	4	095407
2	41	60.3	485	13	09YGS0
3	40	58.8	243	4	095926
4	40	58.8	375	10	09SVA0
5	40	58.8	596	2	09PPR9
6	39	57.4	152	1	029856
7	39	57.4	389	12	09YMM3
8	39	57.4	969	5	023596
9	38	55.9	192	2	09ZLB0
10	38	55.9	363	2	09KPN4
11	38	55.9	2124	10	023057
12	37	54.4	117	4	09UFL7
13	37	54.4	162	2	09X6D7
14	37	54.4	250	2	09XAB0
15	37	54.4	337	2	09K4A8
16	37	54.4	385	5	09NGP7
17	37	54.4	429	4	09NW70
18	37	54.4	462	10	09ZRQ1
19	37	54.4	535	5	097282

20	37	54.4	557	11	09QZF2
21	37	54.4	566	10	09S736
22	37	54.4	600	5	097334
23	37	54.4	687	5	09NH94
24	37	54.4	1943	5	062341
25	36.5	53.7	515	5	016968
26	36	52.9	234	2	09RTT4
27	36	52.9	302	12	084693
28	36	52.9	521	2	025433
29	36	52.9	546	2	09RTX6
30	36	52.9	615	11	09JU98
31	36	52.9	623	11	062510
32	36	52.9	689	1	026380
33	36	52.9	756	4	09NRF3
34	36	52.9	1051	3	001694
35	36	52.9	9477	2	09L4X3
36	35.5	52.2	322	10	09SPK6
37	35	51.5	56	12	098303
38	35	51.5	101	4	09V4H3
39	35	51.5	148	6	09N018
40	35	51.5	177	2	09LAZ4
41	35	51.5	217	12	073551
42	35	51.5	224	2	054370
43	35	51.5	242	5	09VQA4
44	35	51.5	244	2	09S256
45	35	51.5	256	5	044520

ALIGNMENTS

RESULT	1
095407	
ID	095407
AC	095407: PRELIMINARY; PRT; 300 AA.
DT	01-MAY-1999 (TREMBlrel. 10, Created)
DT	01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT	01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE	DECOY RECEPTOR 3 (M68) (M68C) (M68E).
GN	DCR3 OR TR6.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=99087326; PubMed=9872321;
RA	Pitti R.M., Marsters S.A., Lawrence D.A., Roy M., Kischkel F.C.,
RA	Dowd P., Huang A., Donahue C.J., Sherwood S.W., Baldwin D.T.,
RA	Godowski P.J., Wood W.I., Gurney A.L., Hillan K.J., Cohen R.L.,
RA	Goddard A.D., Botstein D., Ashkenazi A.;
RT	"Genomic amplification of a decoy receptor for Fas ligand in lung and
RT	colon cancer.";
RL	Nature 396:699-703(1998).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	TISSUE=BLOOD;
RX	MEDLINE=99253915; PubMed=10318773;
RA	Yu K.Y., Kwon B., Ni J., Zhai Y., Ebner R., Kwon B.S.;
RT	"A newly identified member of tumor necrosis factor receptor
RT	superfamily (TR6) suppresses LIGHT-mediated apoptosis.";
RL	J. Biol. Chem. 274:13733-13736(1999).
RN	[3]
RP	SEQUENCE FROM N.A.
RC	TISSUE=PANCREAS;
RX	MEDLINE=20122600; PubMed=10655513;
RA	Bai C., Connolly B., Metzker M.L., Hilliard C.A., Liu X., Sandig V.,
RA	Soderman A., Galloway S.M., Liu Q., Austin C.P., Caskey C.T.;
RT	"Overexpression of M68/pdcr3 in human gastrointestinal tract tumors
RT	independent of gene amplification and its location in a four-gene
RT	cluster.";
RL	Proc. Natl. Acad. Sci. U.S.A. 97:1230-1235(2000).
DR	EMBL; AF104419; AAD03056.1; -.

```

DR EMBL; AF134240; AAD29688.1; -
DR EMBL; AF217796; AAF35244.1; -
DR EMBL; AF217793; AAF33685.1; -
DR EMBL; AF217794; AAF33686.1; -
DR HSSP; P25942; ICDF.
DR INTERPRO; IPR000561; -
DR INTERPRO; IPR001368; -
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS00030; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -; 1.
DR Receptor.
SQ SEQUENCE 300 AA; 32679 MW; F90AE33718449AF CRC64;

Query Match 100.0%; Score 68; DB 4; Length 300;
Best Local Similarity 100.0%; Pred. No. 0.00041;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPGAECERAVID 13
Db 205 VPGAECERAVID 217

RESULT 2
Q9YGS0 PRELIMINARY; PRT; 485 AA.
AC Q9YGS0;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE IROQUOIS-CLASS HOMEODOMAIN PROTEIN IRX-4 (IROQUOIS-RELATED HOMEBOX
DE TRANSCRIPTION FACTOR).
GN IRX4.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99148101; PubMed=10024241;
RA Bao Z.; Z.; Bruneau B.G.; Seidman J.G.; Seidman C.E.; Cepko C.L.;
RT "Regulation of chamber-specific gene expression in the developing
RT heart by Irx4."
RL Science 283:1161-1164(1999).
CC -1- FUNCTION: REGULATES THE CHAMBER-SPECIFIC EXPRESSION OF MYOSIN
CC ISOFORMS BY ACTIVATING THE EXPRESSION OF THE VENTRICLE MYOSIN
CC HEAVY CHAIN-1 (VMHC1) AND SUPPRESSING THE EXPRESSION OF THE ATRIAL
CC MYOSIN HEAVY CHAIN-1 (AMHC1) IN THE VENTRICLES. MAY PLAY A
CC CRITICAL ROLE IN ESTABLISHING CHAMBER-SPECIFIC GENE EXPRESSION IN
CC THE DEVELOPING HEART.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: VENTRICLES OF THE HEART, DEVELOPING FEATHER
CC BUDS, RETINA, HINDBRAIN.
CC -1- DEVELOPMENTAL STAGE: ALL STAGES OF HEART DEVELOPMENT.
CC -1- SIMILARITY: BELONGS TO THE TALE/IRX FAMILY OF HOMEBOX PROTEINS.
DR EMBL; AF091504; AAD16100.1; -
DR HSSP; P40424; 1B72.
DR INTERPRO; IPR001356; -
DR PFAM; PF00046; homeobox_1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
DR HOMEBOX; DNA-binding; Nuclear protein; Transcription regulation;
KW Activator; Repressor; Developmental protein.
FT DNA_BIND 144 203 HOMEBOX (TALE-TYPE).
SQ SEQUENCE 485 AA; 52773 MW; 81F5A72F9B30C208 CRC64;

Query Match 60.3%; Score 41; DB 13; Length 485;
Best Local Similarity 63.6%; Pred. No. 33;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 VPGAECERAVI 12
Db 318 VPGAECERAVI 328

RESULT 3
O95926 PRELIMINARY; PRT; 243 AA.
AC O95926;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
DE DJ465N24.1 (PUTATIVE NOVEL PROTEIN SIMILAR TO PREDICTED YEAST AND WORM
DE PROTEINS).
GN DJ465N24.1 OR DKFZP564O2082.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wilson S.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA Poustka A.; Klein M.; Mewes H.W.; Gassenhuber J.; Wiemann S.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL031432; CAB37991.1; -
DR EMBL; AL080166; CAB45754.1; -
KW Hypothetical protein.
SQ SEQUENCE 243 AA; 28722 MW; 6F910C13BC09A7F7 CRC64;

Query Match 58.8%; Score 40; DB 4; Length 243;
Best Local Similarity 61.5%; Pred. No. 25;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 VPGAECERAVID 13
Db 173 VPSTEIDRVID 185

RESULT 4
O9SVA0 PRELIMINARY; PRT; 375 AA.
AC O9SVA0;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE HYPOTHETICAL 42.8 KDA PROTEIN.
GN F23K16.210 OR AT4G39580.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M.; Murphy G.; Ridley P.; Hudson S.; Bancroft I.; Mewes H.W.;
RA Mayer K.F.X.; Lemcke K.; Schueller C.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Murphy G.; Ridley P.; Hudson S.; Mewes H.W.; Lemcke K.; Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]

```

RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL078620; CAB44693.1; -;
 DR EMBL; AL161595; CAB80621.1; -;
 DR INTERPRO; IPR001798; -;
 DR INTERPRO; IPR001810; -;
 DR PFAM; PF00646; F-box; 1.
 DR PFAM; PF01344; Kelch; 2.
 KW Hypothetical protein.
 SQ SEQUENCE 375 AA; 42815 MW; 00D89E27170DFE51 CRC64;

Query Match 58.8%; Score 40; DB 10; Length 375;
 Best Local Similarity 70.0%; Pred. No. 38;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 PGAECECERAV 11
 II I III I
 Db 211 PGTCERLTV 220

RESULT 5
 Q9PPR9 PRELIMINARY; PRT; 596 AA.

AC Q9PPR9;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE HYPOTHETICAL PROTEIN U0570.
 GN U0570.
 OS Ureaplasma urealyticum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
 OC Mycoplasmataceae; Ureaplasma.
 OX NCBI_TaxID=2130;
 RN [1]

RP SEQUENCE FROM N.A.
 RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,
 RA Cassell G.H.;
 RT "The complete sequence of Ureaplasma urealyticum: Alternate views of a
 minimal genome and sexually transmitted pathogen."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF002155; AAF30984.1; -;
 DR INTERPRO; IPR000209; -;
 DR INTERPRO; IPR001279; -;
 DR INTERPRO; IPR001587; -;
 DR PFAM; PF00753; lactamase_B; 1.
 DR PFAM; PF02147; UPF0036; 1.
 DR PROSITE; PS00136; SUBTILASE_ASP; UNKNOWN_1.
 KW Hypothetical protein.
 SQ SEQUENCE 596 AA; 66839 MW; A3CFEDB62881E63E CRC64;

Query Match 58.8%; Score 40; DB 2; Length 596;
 Best Local Similarity 61.5%; Pred. No. 60;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 VPGAEECERAV 13
 :II I II III
 Db 202 IPGFSESERVID 214

RESULT 6
 O29856 PRELIMINARY; PRT; 152 AA.

AC O29856;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
 DE HYPOTHETICAL 17.4 KDA PROTEIN.
 GN AF0391.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;

OC Archaeoglobus.
 OX NCBI_TaxID=2234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.P., McDonald L., Utterback T.,
 RA Cotton M.D., Striggs T., Artiach P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon Archaeoglobus fulgidus.";
 RL Nature 390:364-370(1997).
 DR EMBL; AF001077; AAB90848.1; -;
 DR TIGR; AF0391; -;
 KW Hypothetical protein.
 SQ SEQUENCE 152 AA; 17361 MW; 846403FF369ACEA1 CRC64;

Query Match 57.4%; Score 39; DB 1; Length 152;
 Best Local Similarity 72.7%; Pred. No. 23;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPGAEECERAV 11
 II IIIII I
 Db 56 VPDFEECERFV 66

RESULT 7
 Q9YMM3 PRELIMINARY; PRT; 389 AA.
 ID Q9YMM3;
 AC Q9YMM3;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
 DE LDORF-104 PEPTIDE.
 OS Lymantria dispar multicapsid nuclear polyhedrosis virus (LdMNPV).
 OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
 OC Nucleopolyhedrovirus.
 OX NCBI_TaxID=10449;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99124785; PubMed=9887315;
 RA Kuzio J., Pearson M.N., Harwood S.H., Funk C.J., Evans J.T.,
 RA Slavicek J.W., Rohrmann G.F.;
 RT "Sequence and analysis of the genome of a baculovirus pathogenic for
 RT Lymantria dispar."
 RL Virology 253:17-34(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Kuzio J., Pearson M.N., Harwood S.H., Funk C.J., Evans J.T.,
 RA Slavicek J., Rohrmann G.F.;
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Kuzio J.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF081810; AAC70290.1; -;
 SQ SEQUENCE 389 AA; 44514 MW; 7AE08EB9885212D5 CRC64;

Query Match 57.4%; Score 39; DB 12; Length 389;
 Best Local Similarity 58.3%; Pred. No. 59;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 PGAECECERAV 13

gaps 0;

RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Waterston R.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Wilson R., Andrews S.;
RL "The sequence of A. thaliana F5110.";
RN Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Washu;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Waterston R.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA Lamar B., Stoneking T., Stumpf J., Mewes H.W., Lemcke K.,
RA Mayer K.F.x.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF013293; AAB62842.1; -;
DR EMBL; AF195115; AAF02800.1; -;
DR EMBL; AL161472; CAB80854.1; -;
DR INTERPRO; IPR002106; -;
DR PROSITE; PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 2124 AA; 235154 MW; 84892F6BEF3B65D5 CRC64;

Query Match 55.9%; Score 38; DB 10; Length 2124;
Best Local Similarity 66.7%; Pred. No. 4.7e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PGAECECERA 10
DB 783 PGCEDCRRA 791
|||:|:|

RESULT 12
Q9UFL7 PRELIMINARY; PRT; 117 AA.
AC Q9UFL7;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE HYPOTHETICAL 12.7 KDA PROTEIN.
GN DKEZP434K058.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Ottenwaelder B., Obermaier B., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL117583; CAB56005.1; -;
KW Hypothetical protein.
SQ SEQUENCE 117 AA; 12742 MW; 5B2E52156AE7BD29 CRC64;

Query Match 54.4%; Score 37; DB 4; Length 117;
Best Local Similarity 77.8%; Pred. No. 40;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 PGAECECERA 10
DB 87 PGAEHCLRA 95
|||||:|:|

RESULT 13
Q9X6D7 PRELIMINARY; PRT; 162 AA.
ID Q9X6D7;
AC Q9X6D7;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN (EC 4.1.1.39)
DE (FRAGMENT).
GN RBCL.
OS Trichodesmium thiebautii.
OC Bacteria; Cyanobacteria; Oscillatoriales; Trichodesmium.
OX NCBI_TaxID=1208;
RN [1]
RP SEQUENCE FROM N.A.
RA Davies J.T., Wyman M.;
RT "Amplification and characterisation of the rbcL gene from the marine
RL cyanobacterium, Trichodesmium thiebautii.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: RUBISCO CATALYSES TWO REACTIONS: THE CARBOXYLATION OF D-
RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC
CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
CC REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
CC ACTIVE SITE.
CC -!- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + CO(2) -> 3-
PHOSPHO-D-GLYCERATE.
CC -!- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + O(2) -> 3-
PHOSPHO-D-GLYCERATE + 2-PHOSPHOGLYCOLATE.
CC -!- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.
DR EMBL; AF136182; AAD33675.1; -;
DR INTERPRO; IPR000685; -;
DR PFAM; PF00016; RUBISCO_large; 1.
DR PROSITE; PS00157; RUBISCO_LARGE; 1.
KW Photosynthesis; Carbon dioxide fixation; Photorespiration; Lyase;
FT Oxidoreductase; Monooxygenase.
FT NON_TER 1
FT NON_TER 162
FT ACT_SITE 28
SQ SEQUENCE 162 AA; 18102 MW; 37FEE2FID7881358 CRC64;
BINDING OF CO(2) ACTIVATES THE ENZYME.

Query Match 54.4%; Score 37; DB 2; Length 162;
Best Local Similarity 70.0%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 PGAECECERAV 11
DB 71 PTCEECRRV 80
|:|:|:|

RESULT 14
Q9XAB0 PRELIMINARY; PRT; 250 AA.
ID Q9XAB0;
AC Q9XAB0;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
DE HYPOTHETICAL 26.8 KDA PROTEIN.
GN SCF43A.05.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.

Wed Jan 31 13:44:12 2001

RC STRAIN=A3(2);
RA Seeger K., Harris D.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RL the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA James K.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapaites D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RL the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL096837; CAB48892.1; -;
KW Hypothetical protein.
SQ SEQUENCE 250 AA; 26785 MW; 855C13071B6DFBCD CRC64;

Query Match 54.4%; Score 37; DB 2; Length 250;
Best Local Similarity 85.7%; Pred. No. 85;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGAEEC 7
DB 171 LPGAECC 177

RESULT 15
Q9K4A8
ID Q9K4A8 PRELIMINARY; PRT; 337 AA.
AC Q9K4A8;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE CONSERVED HYPOTHETICAL PROTEIN.
GN SC7E4.08C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Seeger K.J., Harris D.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Redenbach M., Kieser H.M., Denapaites D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RL the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL359214; CAB94596.1; -;
SQ SEQUENCE 337 AA; 36763 MW; 80CE1F173CB1D522 CRC64;

Query Match 54.4%; Score 37; DB 2; Length 337;
Best Local Similarity 70.0%; Pred. No. 11e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 VPGAEEC 10

Db 288 VPGTEGCPRA 297

Search completed: January 30, 2001, 16:54:49
Job time: 905 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:47:43 ; Search time 132.77 Seconds
(without alignments)
1.758 Million cell updates/sec

Title: US-09-518-931-2_COPY_205_217

Perfect score: 68

Sequence: 1 VPGAEECERAVID 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA.*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6_COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	68	100.0	300	2	US-08-794-796-2		Sequence 2, Appli
2	45	66.2	437	1	US-07-863-169A-3		Sequence 3, Appli
3	45	66.2	437	2	US-08-429-964-3		Sequence 3, Appli
4	45	66.2	437	3	US-07-935-087-3		Sequence 3, Appli
5	45	66.2	437	4	PCT-US93-08062-3		Sequence 3, Appli
6	45	66.2	444	3	US-07-935-087-7		Sequence 7, Appli
7	37	54.4	229	1	US-08-702-344-9		Sequence 9, Appli
8	34	50.0	377	1	US-08-227-108-17		Sequence 17, Appli
9	34	50.0	377	2	US-09-073-674-17		Sequence 17, Appli
10	34	50.0	424	2	US-08-871-268A-23		Sequence 23, Appli
11	34	50.0	424	3	US-08-871-267B-31		Sequence 31, Appli
12	34	50.0	753	2	US-08-867-941-20		Sequence 20, Appli
13	34	50.0	985	2	US-08-867-941-13		Sequence 13, Appli
14	34	50.0	985	2	US-08-867-941-17		Sequence 17, Appli
15	34	50.0	1000	2	US-08-867-941-12		Sequence 12, Appli
16	34	50.0	1000	2	US-08-867-941-16		Sequence 16, Appli
17	34	50.0	3224	2	US-08-705-660-34		Sequence 34, Appli
18	34	50.0	3224	3	US-08-989-045-34		Sequence 34, Appli
19	33	48.5	192	1	US-08-086-428B-79		Sequence 79, Appli
20	33	48.5	192	2	US-08-468-570-79		Sequence 79, Appli
21	33	48.5	192	2	US-08-290-665A-79		Sequence 79, Appli
22	33	48.5	192	4	PCT-US95-10398-79		Sequence 79, Appli
23	33	48.5	590	1	US-08-448-196A-9		Sequence 9, Appli
24	33	48.5	653	5	5248599-2		Patent No. 5248599
25	33	48.5	1311	1	US-08-340-011-5		Sequence 5, Appli
26	33	48.5	1311	1	US-08-901-710-5		Sequence 5, Appli
27	33	48.5	1338	3	US-08-750-141A-3		Sequence 3, Appli
28	32	47.1	60	3	US-08-476-705A-5		Sequence 5, Appli

29	32	47.1	60	5	5202417-1	Patent No. 5202417
30	32	47.1	65	2	US-08-162-081B-49	Sequence 49, Appli
31	32	47.1	65	2	US-08-780-872-49	Sequence 49, Appli
32	32	47.1	66	3	US-08-782-480-26	Sequence 26, Appli
33	32	47.1	84	5	5202417-2	Patent No. 5202417
34	32	47.1	131	2	US-08-162-081B-41	Sequence 41, Appli
35	32	47.1	131	2	US-08-780-872-41	Sequence 41, Appli
36	32	47.1	192	1	US-08-086-428B-77	Sequence 77, Appli
37	32	47.1	192	2	US-08-468-570-77	Sequence 77, Appli
38	32	47.1	192	2	US-08-290-665A-77	Sequence 77, Appli
39	32	47.1	192	4	PCT-US95-10398-77	Sequence 77, Appli
40	32	47.1	243	3	US-09-191-647-14	Sequence 14, Appli
41	32	47.1	318	1	US-08-537-942A-1	Sequence 1, Appli
42	32	47.1	393	2	US-08-559-303B-76	Sequence 76, Appli
43	32	47.1	395	2	US-08-841-349-9	Sequence 9, Appli
44	32	47.1	471	2	US-08-466-583-2	Sequence 2, Appli
45	32	47.1	471	4	PCT-US95-07820-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-794-796-2
; Sequence 2, Application US/08794796
; Patent No. 5885800
; GENERAL INFORMATION:
; APPLICANT: Emery, John
; APPLICANT: Tan, KB
; APPLICANT: Truneh, Alem
; APPLICANT: Young, Peter
; TITLE OF INVENTION: Tumor Necrosis Related Receptor,
; TITLE OF INVENTION: TR4
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/794,796
; FILING DATE: 04-FEB-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: GH50000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-4026
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 300 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-794-796-2

Query Match 100.0%; Score 68; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 0.00017;

DD 420 VFGFEECEDAV 1300

RESULT 4
US-07-935-087-3
; sequence 3, Application US/079350

US-07-935-087-3
; Sequence 3, Application US/07935087

; Patent No. 6083917
; GENERAL INFORMATION:
; APPLICANT: BROWN, MICHAEL S.
; APPLICANT: GOLDSTEIN, JOSEPH L.
; APPLICANT: REISS, YUVAL
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; TITLE OF INVENTION: THE IDENTIFICATION,
; TITLE OF INVENTION: CHARACTERIZATION,
; TITLE OF INVENTION: AND INHIBITION OF FARNESYL
; TITLE OF INVENTION: PROTEIN TRANSFERASE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: P.O. BOX 4433
; CITY: HOUSTON
; STATE: TEXAS
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1 (converted to ASCII-DOS)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,087
; FILING DATE: 19920824
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/822,011
; FILING DATE: 01/16/92
; ATTORNEY/AGENT INFORMATION:
; NAME: PARKER, DAVID L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTSID:269/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512-320-7200
; TELEFAX: 512-474-7577
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 437 amino acid residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-07-935-087-3

Query Match 66.2%; Score 45; DB 3; Length 437;
Best Local Similarity 81.8%; Pred. No. 2.4;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VPGAECERAV 11
||| |||| ||
DB 420 VPGFEECEDAV 430

RESULT 5

PCT-US93-08062-3
; Sequence 3, Application PC/TUS9308062
; GENERAL INFORMATION:
; APPLICANT:
; SEQUENCE CHARACTERISTICS:
; SEQUENCE CHARACTERISTICS: BROWN, MICHAEL S.
; SEQUENCE CHARACTERISTICS: GOLDSTEIN, JOSEPH L.
; SEQUENCE CHARACTERISTICS: REISS, YUVAL
; SEQUENCE CHARACTERISTICS: MARSTERS, JR., JAMES C.
; ADDRESSEE: METHODS AND COMPOSITIONS FOR
; ADDRESSEE: THE IDENTIFICATION,
; ADDRESSEE: CHARACTERIZATION AND
; ADDRESSEE: INHIBITION OF
; ADDRESSEE: FARNESYLTRANSFERASE
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: P.O. BOX 4433
; CITY: HOUSTON
; STATE: TEXAS
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK/ASKII
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08062
; FILING DATE: AUGUST 24, 1993
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/935,087
; FILING DATE: 24 AUGUST 1992 (24.08.92)
; NAME: UNKNOWN
; ATTORNEY/AGENT INFORMATION:
; NAME: PARKER, DAVID L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTFD377PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512-320-7200
; TELEFAX: 512-474-7577
; TELEX: NOT APPLICABLE
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 437 amino acid residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; PCT-US93-08062-3

Query Match 66.2%; Score 45; DB 4; Length 437;
Best Local Similarity 81.8%; Pred. No. 2.4;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VPGAECERAV 11
||| |||| ||
DB 420 VPGFEECEDAV 430

RESULT 6

US-07-935-087-7
; Sequence 7, Application US/07935087
; Patent No. 6083917
; GENERAL INFORMATION:
; APPLICANT: BROWN, MICHAEL S.
; APPLICANT: GOLDSTEIN, JOSEPH L.
; APPLICANT: REISS, YUVAL
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; TITLE OF INVENTION: THE IDENTIFICATION,
; TITLE OF INVENTION: CHARACTERIZATION,
; TITLE OF INVENTION: AND INHIBITION OF FARNESYL
; TITLE OF INVENTION: PROTEIN TRANSFERASE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: P.O. BOX 4433
; CITY: HOUSTON
; STATE: TEXAS
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1 (converted to ASCII-DOS)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,087

us-09-518-931-2_copy_205_217.rai

Wed Jan 31 13:44:08 2001

```
;
;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-702-344-9

Query Match      54.4%; Score 37; DB 1; Length 229;
Best Local Similarity 66.7%; Pred. No. 30;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 4 AEECERAVI 12
   |||:|:|:|:|
Db 6 AEDCEQAVV 14

RESULT 8
US-08-227-108-17
; Sequence 17, Application US/08227108
; Patent No. 5807726
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/227,108
; FILING DATE: 03-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fanucci, Allan A.
; REGISTRATION NUMBER: 30,256
; REFERENCE/DOCKET NUMBER: 7620-033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 377 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-227-108-17

Query Match      50.0%; Score 34; DB 1; Length 377;
Best Local Similarity 54.5%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 GAEECERAVI 13
   |||:|:|:|:|
Db 223 GTEVCSREVL 233

RESULT 9
US-09-073-674-17
; Sequence 17, Application US/09073674
; Patent No. 5998189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
;
; FILING DATE: 19920824
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: US/07/822,011
; APPLICATION NUMBER: 01/16/92
; FILING DATE: 01/16/92
; ATTORNEY/AGENT INFORMATION:
; NAME: PARKER, DAVID L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTSD:269/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512-320-7200
; TELEFAX: 512-474-7577
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 444 amino acid residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-07-935-087-7

Query Match      66.2%; Score 45; DB 3; Length 444;
Best Local Similarity 81.8%; Pred. No. 2.5;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VPGAEECERAV 11
   ||| ||| ||| ||
Db 420 VPGFEECEDAV 430

RESULT 7
US-08-702-344-9
; Sequence 9, Application US/08702344
; Patent No. 5723315
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: Lavallie, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,344
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWN, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 229 amino acids
; TYPE: amino acid
; STRANDEDNESS:
```

APPLICANT: Benicourt, Claude
APPLICANT: Junien, Jean-Louis
TITLE OF INVENTION: Recombinant Dog Gastric Lipase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Warner-Lambert Company
STREET: 2800 Plymouth Road
CITY: Ann Arbor
STATE: Michigan
COUNTRY: U.S.A.
ZIP: 48105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,674
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Crissey, Todd M.
REGISTRATION NUMBER: 37,807
REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 734 622-7530
TELEFAX: 734 622-1553
TELEX:
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 377 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-073-674-17

Query Match 50.0%; Score 34; DB 2; Length 377;
Best Local Similarity 54.5%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 GAECECAVID 13
| | | | |
Db 223 GTEVCSREVID 233

RESULT 10
US-08-871-268A-23
Sequence 23, Application US/08871268A
Patent No. 5866391
GENERAL INFORMATION:
APPLICANT: Jones, Aubrey R.
TITLE OF INVENTION: Aspergillus Porphobilinogen Synthases
TITLE OF INVENTION: and Nucleic Acids Encoding Same
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5866391o No. 5866391disk of No. 5866391th America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/871,268A
FILING DATE: 09-JUN-1997
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4809.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 424 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5866391e
US-08-871-268A-23

Query Match 50.0%; Score 34; DB 2; Length 424;
Best Local Similarity 75.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 EECERAVI 12
| | | | |
Db 74 EECRAAV 81

RESULT 11
US-08-871-267B-31
Sequence 31, Application US/08871267B
Patent No. 6100057
GENERAL INFORMATION:
APPLICANT: Elrod, Susan L.
APPLICANT: Cherry, Joel R.
APPLICANT: Jones, Aubrey
TITLE OF INVENTION: A Method for Increasing Hemoprotein
TITLE OF INVENTION: Production in Filamentous Fungi
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6100057o No. 6100057disk Of No. 6100057th America, Inc.
STREET: 405 Lexington Avenue - 64th Fl.
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/871,267B
FILING DATE: 9-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rozek, Carol E.
REGISTRATION NUMBER: 36,993
REFERENCE/DOCKET NUMBER: 4771.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-878-9652
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 424 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6100057e
US-08-871-267B-31

Query Match 50.0%; Score 34; DB 3; Length 424;
Best Local Similarity 75.0%; Pred. No. 1.9e+02;

Wed Jan 31 13:44:08 2001

```

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 EECERAVI 12
    ||||| ||:
Db 74 EECRAAV 81

RESULT 12
US-08-867-941-20
; Sequence 20, Application US/08867941
; Patent No. 5977337
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M
; APPLICANT: Du, Run-Pan
; APPLICANT: Wang, Quijun
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H
; TITLE OF INVENTION: LACTOFERRIN RECEPTOR GENES OF MORAXELLA
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/867.941
; FILING DATE: 03-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-681 MIS:jb
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 753 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-867-941-20

Query Match 50.0%; Score 34; DB 2; Length 753;
Best Local Similarity 77.8%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 AEECERAVI 12
    | ||||| |
Db 379 ANECERAPI 387

RESULT 13
US-08-867-941-13
; Sequence 13, Application US/08867941
; Patent No. 5977337
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M
; APPLICANT: Du, Run-Pan
; APPLICANT: Wang, Quijun
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H
; TITLE OF INVENTION: LACTOFERRIN RECEPTOR GENES OF MORAXELLA
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/867.941
; FILING DATE: 03-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-681 MIS:jb
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 753 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-867-941-20

Query Match 50.0%; Score 34; DB 2; Length 753;
Best Local Similarity 77.8%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 AEECERAVI 12
    | ||||| |
Db 379 ANECERAPI 387

RESULT 14
US-08-867-941-17
; Sequence 17, Application US/08867941
; Patent No. 5977337
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M
; APPLICANT: Du, Run-Pan
; APPLICANT: Wang, Quijun
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H
; TITLE OF INVENTION: LACTOFERRIN RECEPTOR GENES OF MORAXELLA
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/867.941
; FILING DATE: 03-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-681 MIS:jb
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 753 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-867-941-20

Query Match 50.0%; Score 34; DB 2; Length 985;
Best Local Similarity 77.8%; Pred. No. 4.6e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 AEECERAVI 12
    | ||||| |
Db 611 ANECERAPI 619

RESULT 15
US-08-867-941-19
; Sequence 19, Application US/08867941
; Patent No. 5977337
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M
; APPLICANT: Du, Run-Pan
; APPLICANT: Wang, Quijun
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H
; TITLE OF INVENTION: LACTOFERRIN RECEPTOR GENES OF MORAXELLA
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/867.941
; FILING DATE: 03-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-681 MIS:jb
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 985 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-867-941-19

```

TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 985 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-867-941-17

Search completed: January 30, 2001, 16:47:44
Job time: 620 sec

Query Match 50.0%; Score 34; DB 2; Length 985;
Best Local Similarity 77.8%; Pred. NO. 4.6e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 AEECERAVI 12
DB 611 ANECERAPI 619

RESULT 15

US-08-867-941-12
; Sequence 12, Application US/08867941
; Patent No. 5977337
; GENERAL INFORMATION:
; APPLICANT: Loomore, Sheena M
; APPLICANT: Du, Run-Pan
; APPLICANT: Wang, Qijun
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H
; TITLE OF INVENTION: LACTOFERRIN RECEPTOR GENES OF MORAXELLA
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/867,941
; FILING DATE: 03-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-681 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1000 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-867-941-12

Query Match 50.0%; Score 34; DB 2; Length 1000;
Best Local Similarity 77.8%; Pred. NO. 4.6e+02;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 AEECERAVI 12
DB 626 ANECERAPI 634

us-09-518-931-2_copy_205_217.ra1

Wed Jan 31 13:44:08 2001

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:45:24 ; Search time 183.67 seconds
(without alignments)
4.840 Million cell updates/sec

Title: US-09-518-931-2_COPY_239_264

Perfect score: 137

Sequence: 1 PEGWGTPRAGAAALQKLRRRLTEL 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_36.*

- 1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT.*
- 2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT.*
- 3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT.*
- 4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT.*
- 5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT.*
- 6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT.*
- 7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT.*
- 8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT.*
- 9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT.*
- 10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT.*
- 11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT.*
- 12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT.*
- 13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT.*
- 14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT.*
- 15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT.*
- 16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT.*
- 17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT.*
- 18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT.*
- 19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT.*
- 20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT.*
- 21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	137	100.0	215	W93585	Human hAPO6 protein
2	137	100.0	271	Y42184	Human mFLINT #1 pr
3	137	100.0	300	W66102	Amino acid sequenc
4	137	100.0	300	W63622	Human tumour necro
5	137	100.0	300	W03099	Human lung TNF-rec
6	137	100.0	300	Y42182	Human FLINT #1 pro
7	137	100.0	300	Y17479	Mammalian tumour n
8	137	100.0	300	Y06817	Human Dcr3 polypep
9	137	100.0	300	W97749	Human tumour necro
10	137	100.0	300	W95082	Orphan receptor (H
11	137	100.0	300	Y77458	Human TNF receptor
12	131	95.6	273	Y42185	Human mFLINT #2 pr

13	131	95.6	302	20	Y42183	Human FLINT #2 pro
14	57	41.6	527	20	Y29515	Human lung tumour
15	53.5	39.1	774	20	W87502	Human N-methyl-D-a
16	53.5	39.1	774	21	Y56128	Human clone NMDA21
17	53.5	39.1	775	15	R66057	Human NMDAR2 recep
18	53.5	39.1	1214	15	R66065	Human NMDAR2 recep
19	53.5	39.1	1214	20	W87509	Human N-methyl-D-a
20	53.5	39.1	1214	21	Y56133	Human N-methyl-D-a
21	53.5	39.1	1219	15	R66063	Human NMDAR2 recep
22	53.5	39.1	1219	20	W87507	Human N-methyl-D-a
23	53.5	39.1	1219	21	Y56131	Human N-methyl-D-a
24	53.5	39.1	1231	15	R66062	Human NMDAR2 recep
25	53.5	39.1	1231	20	W87506	Human N-methyl-D-a
26	53.5	39.1	1231	21	Y56130	Human N-methyl-D-a
27	53.5	39.1	1236	15	R66037	Human N-methyl-D-a
28	53.5	39.1	1236	20	W85574	Human N-methyl-D-a
29	53.5	39.1	1236	21	Y56111	Human N-methyl-D-a
30	53.5	39.1	1239	15	R66064	Human NMDAR2 recep
31	53.5	39.1	1239	20	W87508	Human N-methyl-D-a
32	53.5	39.1	1239	21	Y56132	Human N-methyl-D-a
33	53.5	39.1	1244	15	R66061	Human NMDAR2 recep
34	53.5	39.1	1244	20	W87505	Human N-methyl-D-a
35	53.5	39.1	1244	21	Y56129	Human N-methyl-D-a
36	50	36.5	67	19	W56734	Nucellus specific
37	50	36.5	797	18	W35577	TNF-RI-DD ligand p
38	46.5	33.9	142	20	Y13138	Human secreted pro
39	46	33.6	785	19	W72195	HSV-2 strain SB5 C
40	45.5	33.2	149	19	W62901	Mutant of the firs
41	45.5	33.2	149	19	W51846	Amino acid sequenc
42	45.5	33.2	149	19	W62902	Mutant of the firs
43	45.5	33.2	247	20	W88359	Human lymphocyte a
44	45.5	33.2	338	20	W88361	Human lymphocyte a
45	45.5	33.2	380	12	R13272	Polyhedrin-soluble

ALIGNMENTS

RESULT 1	
W93585	
ID	W93585 standard; Protein; 215 AA.
XX	
AC	W93585;
XX	
DT	18-JUN-1999 (first entry)
XX	
DE	Human hAPO6 protein.
XX	
KW	Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
KW	developmental abnormality; gestational abnormality; prostate cancer;
KW	AP06; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
KW	cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
KW	apoptosis; human.
OS	Homo sapiens.
XX	
PN	W09911791-A2.
XX	
PD	11-MAR-1999.
XX	
PF	04-SEP-1998; 98WO-US18393.
XX	
PR	05-SEP-1997; 97US-0924634.
XX	
PA	(UNIW) UNIV WASHINGTON.
XX	
PI	Chaudhary PM;
XX	
DR	WPI; 1999-205191/17.
DR	N-PSDB; X23419.
XX	
PT	New Tumor Necrosis Factor family receptor polypeptides and ligands -
PT	useful for diagnosis and treatment of prostate cancer and

Wed Jan 31 13:44:13 2001

developmental or gestational abnormalities

Claim 29; Fig 9; 156pp; English.

This invention describes isolated Tumor Necrosis Factor (TNF) family receptor polypeptides: APO4, APO6, APO8 and APO9 or their active fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or their active fragments. APO4 is useful for diagnosing prostate cancer by determining levels of APO4 in an individual. Prostate cancer can also be treated using APO4 selective binding agents linked to a therapeutic moiety. APO4 polypeptides are also useful for identifying selective binding agents, useful in diagnosis/treatment of disease by binding of agents to the polypeptide/active fragment which is extracellular, or expressed on the cell surface. The binding is preferably performed in vivo. APO4 polypeptides/active fragments are also useful for screening for agonists and antagonists by binding and observing the change in APO4 activity. Effective pharmacological agents useful in diagnosis or treatment of disease are also identified using APO4 polypeptides/active fragments and APO4 signal transducer molecules that specifically interact with a cytoplasmic domain of APO4 and detecting a change in level of APO4 activity. The method is performed in vivo or in vitro. APO polypeptides are all useful as immunogens for preparing antibodies. APO4 is also useful for diagnosis/treatment of developmental or gestational abnormalities. APO8 was transfected to human breast carcinoma cell line MCF-7, and induced apoptosis.

Sequence 215 AA;

Query Match 100.0%; Score 137; DB 20; Length 215;
 Best Local Similarity 100.0%; Pred. No. 3.9e-12;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEGWGTPPRAGRAALQKLRLRLTEL 26
 Db 154 pegwgtppragaaqlqlrrrltel 179

RESULT 2

Y42184
 ID Y42184 standard; Protein; 271 AA.

XX Y42184;

XX 17-DEC-1999 (first entry)

XX Human mFLINT #1 protein sequence;

XX Human: FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
 KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
 KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
 KW reperfusion-associated injury; aplastic anaemia; differentiation;
 KW growth; myelodysplastic syndrome; pancytopenic condition;
 KW myocardial ischaemia.

XX Homo sapiens.

XX WO9950413-A2.

XX 07-OCT-1999.

XX 30-MAR-1999; 99WO-US06797.

XX 30-MAR-1998; 98US-0079856.

XX 20-MAY-1998; 98US-0086074.

XX 09-SEP-1998; 98US-0099643.

XX 17-DEC-1998; 98US-0112577.

XX 18-DEC-1998; 98US-0112703.

XX 18-DEC-1998; 98US-0112933.

XX 22-DEC-1998; 98US-0113407.

XX (ELIL) LILLY & CO ELI.

PI Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
 PI Hui KY, Kharitonkov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
 PI Song HY, Wang J, Wu X, Zuckerman SH;

XX WPI; 1999-591319/50.

DR N-PSDB; Z25377.

XX Use of mature FLINT for treating acute liver failure, inflammation,
 PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
 PT and proinflammatory activity

PS Claim 31; Fig 3; 99pp; English.

XX The present invention describes therapeutic applications of mature FLINT
 CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
 CC superfamily, which is a member of the tumour necrosis factor receptor
 CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
 CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
 CC injury or disorder such as hypercoagulation (including use with
 CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
 CC or disorder, Type I diabetes, cancer, cell damage or damage to an
 CC innocent bystander tissue that is induced by a chemotherapeutic agent or
 CC therapeutic irradiation, treating haematopoietic progenitor cells that
 CC have been exposed to therapeutic radiation or chemotherapy, aplastic
 CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
 CC also used for promoting the growth or differentiation of a haematopoietic
 CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
 CC resulting from abnormal myocardial ischaemia. The present sequence
 CC represents human mFLINT.

XX Sequence 271 AA;

Query Match 100.0%; Score 137; DB 20; Length 271;

Best Local Similarity 100.0%; Pred. No. 5e-12;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEGWGTPPRAGRAALQKLRLRLTEL 26

Db 210 pegwgtppragaaqlqlrrrltel 235
 |||||

RESULT 3

W66102
 ID W66102 standard; Protein; 300 AA.

XX W66102;

XX 02-DEC-1998 (first entry)

XX Amino acid sequence of tumour necrosis related receptor (TR4).

XX Human; tumour necrosis related receptor; TR4; agonist; antagonist;
 KW inhibition; chronic; acute; inflammation; arthritis; septicaemia;
 KW autoimmune disease; transplant rejection; stroke; cancer;
 KW Alzheimer's disease.

XX Homo sapiens.

XX EP861850-A1.

XX 02-SEP-1998.

XX 20-JAN-1998; 98EP-0300382.

XX 04-FEB-1997; 97US-0794796.

XX (SMTK) SMITHKLINE BEECHAM CORP.

XX Emery J, Tan KB, Truneh A, Young PR;

XX WPI; 1998-508248/44.

XX DR

DR N-PSDB; V07654.

XX New DNA encoding tumour necrosis related receptor - used to treat

PT and prevent e.g. inflammation, arthritis, septicaemia, autoimmune

PT diseases, transplant rejection, infection, stroke, ischaemia, AKDS,

PT stenosis, AIDS, bone disorders and cancer

XX

PS Claim 1; Fig 1; 21pp; English.

XX This is the amino acid sequence of the human tumour necrosis related

CC receptor (TR4), used in the method of the invention. The TR4 protein

CC or its agonist can be used to treat a subject in need of enhanced

CC TR4 polypeptide activity. The antagonist is used to inhibit TR4

CC polypeptide activity. The active agents can be used for the

CC treatment and prevention of diseases such as chronic and acute

CC inflammation, arthritis, septicaemia, autoimmune diseases, transplant

CC rejection, stroke, cancer, Alzheimer's disease.

XX

SQ Sequence 300 AA;

Query Match 100.0%; Score 137; DB 19; Length 300;

Best Local Similarity 100.0%; Pred. No. 5.5e-12;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEGWGTPPRAGRAALQLKLRRLTEL 26

Db 239 pegwgtppragaaqlklrrrltel 264

|||||

RESULT 4

W63622

ID W63622 standard; Protein; 300 AA.

XX

AC W63622;

XX

DT 26-OCT-1998 (first entry)

XX

DE Human tumour necrosis factor receptor-6 alpha protein.

XX

XX Human tumour necrosis factor receptor-6 alpha; TNFR-6 beta;

XX Human tumour necrosis factor receptor-6 alpha; TNFR-6 beta;

KW endothelial cells; keratinocytes; normal prostate; apoptosis;

KW prostate tumour tissue.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Peptide 1..300

FT Protein 31..300

FT /note= "TNFR-6 alpha"

FT Region 31..282

FT /note= "Soluble extracellular domain"

XX

PN W09830694-A2.

XX

PD 16-JUL-1998.

XX

PF 13-JAN-1998; 98WO-US00153.

XX

PR 14-JAN-1997; 97US-0035496.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Ebner R, Peng P, Gentz RL, Ni J, Ruben SM, Yu G;

XX

DR WPI; 1998-399142/34.

DR N-PSDB; V39085.

XX

PT Human tumour necrosis factor receptors 6-alpha and 6-beta - used in

PT the diagnosis of immune system-related disorder(s)

XX

PS Claim 20; Fig 1; 91pp; English.

XX

CC The present sequence represents the human tumour necrosis factor

CC receptor-6 alpha (TNFR-6 alpha) protein. The invention also provides

CC for the TNFR-6 beta protein (W63623). TNFR-6 alpha and TNFR-6 beta are

CC members of the tumour necrosis factor receptor (TNFR) family. TNFRs

CC are expressed in endothelial cells, keratinocytes, normal prostate and

CC prostate tumour tissue. For a number of disorders of these cells,

CC particularly of the immune system, substantially altered (whether

CC increased or decreased) levels of TNFR-6 alpha and/or TNFR-6 beta gene

CC expression can be detected, therefore the TNFR-6 alpha and TNFR-6 beta

CC polypeptides, nucleic acids and antibodies are claimed to be useful in

CC the diagnosis of such disorders. Mutations of the TNFR-6 alpha and

CC TNFR-6 beta genes can also be detected. The TNFR polypeptides are

CC also claimed to be useful for identifying ligands which may be useful

CC in the treatment of apoptosis related disorders.

XX

SQ Sequence 300 AA;

Query Match 100.0%; Score 137; DB 19; Length 300;

Best Local Similarity 100.0%; Pred. No. 5.5e-12;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEGWGTPPRAGRAALQLKLRRLTEL 26

Db 239 pegwgtppragaaqlklrrrltel 264

|||||

RESULT 5

Y03099

ID Y03099 standard; Protein; 300 AA.

XX

AC Y03099;

XX

DT 09-DEC-1999 (first entry)

XX

DE Human lung TNF-receptor protein.

XX

KW Tumour necrosis factor; TNF; TNF receptor; human; lung; gene therapy;

KW detection; immunoassay; diagnosis; disease; immune system; tumour;

KW osteogenic system; cardiovascular system; central nervous system; asthma;

KW peripheral nervous systems; transplant incompatibility; antitumor;

KW rheumatoid arthritis; antiasthmatic; antiarthritic.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 134..1036

FT /*tag= a

FT /product= "TNF-receptor"

XX

PN DE19809978-A1.

XX

PD 16-SEP-1999.

XX

PF 09-MAR-1998; 98DE-1009978.

XX

PR 09-MAR-1998; 98DE-1009978.

XX

PA (BADI) BASF AG.

XX

PI Kroeger B;

XX

DR WPI; 1999-519473/44.

DR N-PSDB; Z09998.

XX

PT New soluble member of tumor necrosis factor receptor family, useful for

PT identification specific modulators and for treating disease e.g. tumors

PT

XX Claim 1; Page 8-9; 10pp; German.

PS

XX This invention describes a novel tumour necrosis factor (TNF) receptor

CC (I) isolated from human lung tissue. (I) is used: (i) to raise specific

PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
PT and proinflammatory activity
XX Claim 30; Fig 1; 99pp; English.
PS
XX The present invention describes therapeutic applications of mature FLINT
CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
CC (mFLINT), which is a member of the tumour necrosis factor receptor
CC superfamily, is used for treating acute liver failure, inflammation of
CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
CC injury or disorder such as hypercoagulation (including use with
CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
CC or disorder, Type I diabetes, cancer, cell damage or damage to an
CC innocent bystander tissue that is induced by a chemotherapeutic agent or
CC therapeutic irradiation, treating haematopoietic progenitor cells that
CC have been exposed to therapeutic radiation or chemotherapy, aplastic
CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
CC also used for promoting the growth or differentiation of a haematopoietic
CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
CC resulting from abnormal myocardial ischaemia. The present sequence
CC represents human FLINT.
XX Sequence 300 AA;
SQ

Query Match 100.0%; Score 137; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 5.5e-12;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEGWGTPPRAGRAALQKLRRRLTEL 26
Db 239 pegwgtpragraalqlkrrrltel 264
|||||

RESULT 7
Y17479
ID Y17479 standard; Protein; 300 AA.
XX
AC Y17479;
XX
DT 02-AUG-1999 (first entry)
XX
DE Mammalian tumour necrosis factor receptor OPG-2.
XX
XX Tumour necrosis factor receptor; TNF receptor; OPG-2; Paget's disease;
KW osteopenic disorder; osteoclast activity; primary osteoporosis;
KW hyperglycaemia; osteolytic metastasis; immune response; cancer.
XX
XX Mammalia.
OS
XX
XX WO9926977-A1.
PN
XX
XX 03-JUN-1999.
PD
XX
XX 24-NOV-1998; 98WO-US25065.
PF
XX
XX 17-FEB-1998; 98US-0074896.
PR
XX 24-NOV-1997; 97US-0066446.
PR
XX
XX (BIOJ) BIOGEN INC.
PA
XX
XX Tschopp J;
PI
XX
XX WPI; 1999-347693/29.
DR
XX N-PSDB; X76052.
DR
XX New tumour necrosis factor family receptor OPG-2
PT
XX
XX Claim 1; Page 18; 22pp; English.
PS
XX
XX The present sequence represents a mammalian tumour necrosis factor
CC receptor, designated OPG-2. OPG-2, is a member of the tumour necrosis
CC

CC antibodies (Ab); (ii) to screen for specific (antagonists or ligands
CC (A), potential therapeutic agents; and (iii) therapeutically (optionally
CC expressed from a gene therapy vector) in conditions associated with a
CC deficit of (I). Ab are used: (a) for qualitative or quantitative
CC detection of (I) in standard immunoassays (for diagnosis of disease, or
CC susceptibility, or for monitoring); and (b) as therapeutic inhibitors in
CC cases where (I) is overexpressed. Nucleic acid (II) that encodes (I) is
CC used: (A) for recombinant production of (I); (B) also its oligonucleotide
CC fragments, in standard hybridization and/or amplification assays; (C) as
CC source of antisense molecules or ribozymes; and (D) to produce transgenic
CC animals (for studying (patho)physiology of (I)). Diseases possibly
CC associated with under- or over-expression of (I) are those of the immune,
CC osteogenic, cardiovascular and central or peripheral nervous systems,
CC tumors, transplant incompatibility, asthma and rheumatoid arthritis. The
CC products of the invention have antitumor, antiasthmatic and
CC antiarthritic activity. This sequence represents the TNF-receptor of the
CC invention.
XX Sequence 300 AA;
SQ

Query Match 100.0%; Score 137; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 5.5e-12;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEGWGTPPRAGRAALQKLRRRLTEL 26
Db 239 pegwgtpragraalqlkrrrltel 264
|||||

RESULT 6
Y42182
ID Y42182 standard; Protein; 300 AA.
XX
AC Y42182;
XX
DT 17-DEC-1999 (first entry)
XX
DE Human FLINT #1 protein sequence.
XX
XX Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
KW reperfusion-associated injury; aplastic anaemia; differentiation;
KW growth; myelodysplastic syndrome; pancytopenic condition;
KW myocardial ischaemia.
XX
XX Homo sapiens.
OS
XX
XX WO9950413-A2.
PN
XX
XX 07-OCT-1999.
PD
XX
XX 30-MAR-1999; 99WO-US06797.
PF
XX
XX 30-MAR-1998; 98US-0079856.
PR
XX 20-MAY-1998; 98US-0086074.
PR
XX 09-SEP-1998; 98US-0099643.
PR
XX 17-DEC-1998; 98US-0112577.
PR
XX 18-DEC-1998; 98US-0112703.
PR
XX 18-DEC-1998; 98US-0112933.
PR
XX 22-DEC-1998; 98US-0113407.
PR
XX
XX (ELIL) LILLY & CO ELI.
PA
XX
XX Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
PI Hui KY, Kharitonkov A, Mizrahi J, Na S,
PI Song HY, Wang J, Wu X, Zuckerman SH;
XX
XX WPI: 1999-591319/50.
DR
XX N-PSDB; Z25375.
DR
XX
XX Use of mature FLINT for treating acute liver failure, inflammation,
PT

CC factor receptor family, and can be used: (i) to raise specific
 CC antibodies (Ab), (ii) to treat osteopenic disorders associated with
 CC excessive osteoclast activity, e.g. primary osteoporosis, Paget's
 CC disease, hyperglycaemia of malignancy, or osteolytic metastases; (iii)
 CC for affinity purification of cognate ligands, and (iv) to screen for
 CC ligands (antagonists or agonists). Ab, or other OPG-2 blocking agents
 CC such as soluble forms of the protein, are used to prevent, or reduce
 CC severity of, an immune response, and for treating cancer. They can also
 CC be used in diagnostic assays. The nucleic acid sequence encoding OPG-2
 CC can be used as a probe to isolate related sequences from other species.
 XX
 XX

SQ Sequence 300 AA;

Query Match 100.0%; Score 137; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 5.5e-12;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PEGWGPTPRAGRAALQLKLRRLTEL 26
 |||||
 DB 239 pegwgtpragaaqlklrrltel 264

RESULT 8

ID Y06817 standard; Protein; 300 AA.

XX AC Y06817;

XX DT 24-JUN-1999 (first entry)

XX DE Human Dcr3 polypeptide.

XX KW Dcr3 polypeptide; tumour necrosis factor receptor; TNFR; Fas ligand;
 KW apoptosis; T cell mediated immune response; allergy; asthma; cancer;
 KW rheumatoid arthritis; Crohn's disease; guest vs. host disease; human;
 KW gene therapy.

XX OS Homo sapiens.

XX PN WO9914330-A1.

XX PD 25-MAR-1999.

XX PF 18-SEP-1998; 98WO-US19661.

XX PR 30-JUL-1998; 98US-0094640.

XX PR 18-SEP-1997; 97US-0059288.

XX PA (GETH) GENENTECH INC.

XX PI Ashkenazi AJ, Botstein D, Dodge KH, Goddard A, Gurney AL;
 PI Kim KJ, Lawrence DA, Pitti R, Roy MA, Tumas DB;
 PI Wood WI;

XX WPI: 1999-244032/20.

XX DR N-PSDB; X32744.

XX PT Dcr3 polypeptide related to tumor necrosis factor receptor
 XX
 XX

XX PS Claim 5; Fig 1; 88pp; English.

XX This represents a human Dcr3 polypeptide, a homologue of tumour necrosis
 CC factor receptor (TNFR) polypeptide. Host cells containing a vector
 CC comprising the Dcr3 nucleic acid can be used for the recombinant
 CC expression of the protein. Dcr3 binds to Fas ligand, so it (or its
 CC chimeras) are useful for modulating apoptosis in mammalian cells, also
 CC other Fas-ligand induced activities, particularly to inhibit T cell
 CC mediated immune responses, e.g. in treatment of allergy, asthma,
 CC rheumatoid arthritis, Crohn's disease, guest vs. host disease etc. Dcr3
 CC may also be used to identify specific binding proteins, potential
 CC inhibitors. Antibodies against Dcr3 are used to treat cancer.
 CC Specifically of the lung and colon, also in diagnosis and for affinity

CC purification of the protein. Detecting mutations in the gene for Dcr3 is
 CC also used to diagnose cancer, or predisposition to it. Dcr3 nucleic acid
 CC is useful as hybridization probe to detect genomic or related sequences;
 CC for chromosome and gene mapping; as source of antisense sequences; for
 CC expression of recombinant Dcr3 and to generate transgenic animals (for
 CC development and screening of therapeutic agents), also for in vivo or
 CC ex vivo gene therapy.
 XX
 XX

SQ Sequence 300 AA;

Query Match 100.0%; Score 137; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 5.5e-12;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PEGWGPTPRAGRAALQLKLRRLTEL 26
 |||||
 DB 239 pegwgtpragaaqlklrrltel 264

RESULT 9

W97749

ID W97749 standard; Protein; 300 AA.

XX AC W97749;

XX DT 21-MAY-1999 (first entry)

XX DE Human tumour necrosis factor receptor ZTNFR-5.

XX KW ZTNFR-5; tumour necrosis factor receptor; TNFR; human;
 KW cell maturation; bone cell regulation.
 XX

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
 FT Peptide 1..23
 FT Protein /note= "signal peptide"
 FT Domain 24..300
 FT Domain /note= "mature protein"
 FT Region 24..194
 FT Region /note= "extracellular domain"
 FT Region 49..71
 FT Region /note= "cysteine-rich pseudo-repeat 1"
 FT Region 72..113
 FT Region /note= "cysteine-rich pseudo-repeat 1"
 FT Region 114..151
 FT Region /note= "cysteine-rich pseudo-repeat 1"
 FT Region 152..194
 FT Region /note= "cysteine-rich pseudo-repeat 1"
 XX

XX PN WO9904001-A1.

XX PD 28-JAN-1999.

XX PF 21-JUL-1998; 98WO-US15072.

XX PR 21-JUL-1997; 97US-0053203.

XX PA (ZYMO) ZYMOGENETICS INC.

XX PI Farrah TM;

XX WPI: 1999-132245/11.

XX DR N-PSDB; X07226.

XX PT Novel tumour necrosis factor receptor ZTNFR5 - useful for
 XX regulating maturation of TNF-ligand bearing cells
 XX
 XX

XX PS Claim 1; Page 84-85; 109pp; English.

XX This polypeptide comprises a new, secreted tumour necrosis factor
 CC receptor (see W97749), designated ZTNFR-5. Novel ZTNFR-5 encoding

CC bone mass and muscle metabolism. HUMAN NTR-1 receptors are also useful
CC for screening for novel binding agents, and cognate ligands, which may be
CC used to treat disorders associated with HUMAN NTR-1 imbalance.

XX Sequence 300 AA;

Query Match 100.0%; Score 137; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 5.5e-12;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PEGWGPTPRAGRAALQLKRRRLTEL 26
Db 239 pegwgptpragaaqlklrrrltel 264

RESULT 11

Y77458
ID Y77458 standard; Protein; 300 AA.

XX Y77458;

XX 05-JUN-2000 (first entry)

XX Human TNF receptor-like protein, HDTEA84.

XX TNF receptor family; tumour necrosis factor; HDTEA84; HSLJD378;
KW Rank-like protein; RANKL; immune disorder; inflammation; allergy;
KW immunosuppressant; antiarthritic; antirheumatoid; antiinflammatory;
KW dermatological; antithyroid;

XX Homo sapiens.

OS WO200001817-A2.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-US12366;

XX 06-JUL-1998; 98US-0110938;

XX 13-JUL-1998; 98US-0114466;

XX 23-JUL-1998; 98US-0093897;

XX 12-AUG-1998; 98US-0132968;

XX 18-AUG-1998; 98US-0136214;

XX 11-SEP-1998; 98US-0099999;

XX (SCHE) SCHERING CORP.

XX Bates EEM, Lebecque SJE, Murphy EE, Mattson JD, Gorman DM;

XX Hedrick JA, Wang L, Zlotnik A, Murgolo NJ, Greene JR, Johnston JA;

XX Bazan JF, Mahony D, Lees EM;

XX WPI; 2000-171015/15.

XX N-PSDB; 792404.

XX New isolated mammalian genes, used to develop products for treating

XX e.g. immune, inflammatory or allergic abnormalities, cancers or

XX degenerative conditions

XX Claim 24; Page 157; 218pp; English.

XX The invention relates to a number of primate and/or rodent proteins, and
XX the genes which encode them. The invention encompasses human dendritic
XX cell prostaglandin transporter (PC-EGT); the TNF (tumour necrosis
XX factor) receptor family-related proteins HDTEA84, HSLJD378 and RANKL;
XX human CC chemokine HCC5; human deubiquitinating proteins DUB1 and DUB
XX 12; human MD-1 and human and murine MD-2 proteins, which exhibit the
XX properties of ligands for proteins comprising a leucine-rich motif (LRR);
XX human cyclin E2; cDNAs encoding these proteins; and antibodies against
XX these proteins. The proteins can be used for modulating the physiology or
XX development of a cell. They can be used to mediate uptake of substrates
XX (e.g., prostaglandin-like molecules), to modulate or mediate cellular
XX interactions (e.g., induce or prevent trafficking, proliferation, or

CC polynucleotides and polypeptides were initially identified by
CC querying an expressed sequence tag (EST) database for sequences
CC homologous to conserved motifs within the TNF receptor family.

CC Based on this search, a contig of 16 ESTs (see X07226) was
CC constructed. ZTNFR-5 polypeptides comprise 4 cysteine-rich repeats
CC (see also W97750-55) that are homologous to other TNF receptors, in
CC particular the soluble, secreted TNF receptor osteoprotegerin. The
CC ZTNFR-5 polypeptide can be prepared by recombinant methods. can be used to
CC polypeptide, especially the extracellular domain, can be used to
CC generate a soluble variant of ZTNFR-5. The polypeptides and
CC nucleic acids can be used to screen for ligands, agonists and
CC antagonists of ZTNFR-5. The polypeptides can be used in bone cell
CC regulation and to regulate the maturation of TNF ligand-bearing
CC cells such as T- or B-cells, lymphocytes, peripheral blood
CC mononuclear cells, polymorphonuclear leukocytes, fibroblasts or
CC haematopoietic cells.

XX Sequence 300 AA;

Query Match 100.0%; Score 137; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 5.5e-12;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PEGWGPTPRAGRAALQLKRRRLTEL 26
Db 239 pegwgptpragaaqlklrrrltel 264

RESULT 10

W95082
ID W95082 standard; Protein; 300 AA.

XX W95082;

XX 20-MAY-1999 (first entry)

XX Orphan receptor (HUMAN NTR-1) polypeptide.

XX HUMAN NTR-1; orphan receptor; osteoprotegerin; OPG; TNFR; human;
XX tumour necrosis factor receptor; muscle disorder; bone mass; screening;
XX muscle metabolism; binding agent; cognate ligand.

XX Homo sapiens.

OS WO9907738-A2.

XX 18-FEB-1999.

XX 04-AUG-1998; 98WO-US16202.

XX 06-AUG-1997; 97US-0054869.

XX (PROC) PROCTER & GAMBLE CO.

XX (REG-) REGENERON PHARM INC.

XX Masiakowski PJ, Morris J, Valenzuela DM;

XX WPI; 1999-167365/14.

XX N-PSDB; X22300.

XX Novel orphan human receptor polypeptide and nucleic acid - useful as

XX diagnostic reagents and for treatment of muscle disorders

XX Claim 7; Page 21; 23pp; English.

XX This represents a HUMAN NTR-1 polypeptide, a novel orphan receptor. The
XX protein is related to osteoprotegerin (OPG) and to tumour necrosis factor
XX receptor (TNFR). Host cells transfected with a vector comprising the
XX HUMAN NTR-1 nucleic acid are used for the recombinant expression of the
XX protein. HUMAN NTR-1 proteins and antibodies immuno specific for the
XX protein are useful for diagnosis and treatment of humans and animals,
XX especially muscle disorders, as the receptor is involved in regulation of

CC differentiation of cells), or are intracellular proteins which are
 CC important in various cellular processes such as the deubiquitination of
 CC proteins or cell cycle regulation. The products can be used for
 CC treating medical conditions such as immune, inflammatory or allergic
 CC disorders, or abnormal cellular proliferation, for example, cancers or
 CC degenerative conditions. They can be used to modulate immune responses in
 CC disease states e.g., autoimmune disorders, including rheumatoid
 CC arthritis, systemic lupus erythematosus, Hashimoto's autoimmune
 CC thyroiditis, as well as acute and chronic inflammatory responses in which
 CC T cell activation, expansion, and/or immunological T cell memory play an
 CC important role. Sequences Y77458-Y77461 and Y77465-Y77468 represent
 CC TNF receptor family-related proteins. Y77458 is the human protein
 CC HDPEA84, Y77459-Y77461 are human HSLJD37R proteins, Y77465 is
 CC murine Rank-like protein RANKL, and Y77466-Y77468 are human RANKL
 CC proteins.

XX Sequence 300 AA;

Query Match 100.0%; Score 137; DB 21; Length 300;
 Best Local Similarity 100.0%; Pred. No. 5.5e-12;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEGWGTPPRAGRAALQLKLRRLTEL 26
 Y42185
 ID Y42185 standard; Protein: 273 AA.
 XX AC Y42185;
 XX DT 17-DEC-1999 (first entry)
 XX DE Human mFLINT #2 protein sequence.

XX Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
 KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
 KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
 KW reperfusion-associated injury; aplastic anaemia; differentiation;
 KW growth; myelodysplastic syndrome; pancytopenic condition;
 KW myocardial ischaemia.

XX Homo sapiens.

XX WO9950413-A2.

XX PD 07-OCT-1999.

XX PF 30-MAR-1999; 99WO-US06797.

XX PR 30-MAR-1998; 98US-0079856.

XX PR 20-MAY-1998; 98US-0086074.

XX PR 09-SEP-1998; 98US-0099643.

XX PR 17-DEC-1998; 98US-0112577.

XX PR 18-DEC-1998; 98US-0112703.

XX PR 22-DEC-1998; 98US-0112933.

XX PR 22-DEC-1998; 98US-0113407.

XX PA (ELIL) LILLY & CO ELI.

XX PI Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;

XX PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;

XX PI Song HY, Wang J, Wu X, Zuckerman SH;

XX DR WPI: 1999-591319/50.

XX DR N-PSDB; 225378.

XX

PS Example 2; Fig 4; 99pp; English.

XX The present invention describes therapeutic applications of mature FLINT
 CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
 CC (mFLINT), which is a member of the tumour necrosis factor receptor
 CC superfamily, is used for treating acute liver failure, inflammation of
 CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
 CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
 CC injury or disorder such as hypercoagulation (including use with
 CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
 CC or disorder, Type I diabetes, cancer, cell damage or damage to an
 CC innocent bystander tissue that is induced by a chemotherapeutic agent or
 CC therapeutic irradiation, treating haematopoietic progenitor cells that
 CC have been exposed to therapeutic radiation or chemotherapy, aplastic
 CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
 CC also used for promoting the growth or differentiation of a haematopoietic
 CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
 CC resulting from abnormal myocardial ischaemia. The present sequence
 CC represents human mFLINT.

XX Sequence 273 AA;

Query Match 95.6%; Score 131; DB 20; Length 273;
 Best Local Similarity 96.2%; Pred. No. 3.6e-11;
 Matches 25; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PEGWGTPPRAGRAALQLKLRRLTEL 26
 Y42183
 ID Y42183 standard; Protein: 302 AA.
 XX AC Y42183;
 XX DT 17-DEC-1999 (first entry)
 XX DE Human FLINT #2 protein sequence.

XX Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
 KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
 KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
 KW reperfusion-associated injury; aplastic anaemia; differentiation;
 KW growth; myelodysplastic syndrome; pancytopenic condition;
 KW myocardial ischaemia.

XX Homo sapiens.

XX WO9950413-A2.

XX PD 07-OCT-1999.

XX PF 30-MAR-1999; 99WO-US06797.

XX PR 30-MAR-1998; 98US-0079856.

XX PR 20-MAY-1998; 98US-0086074.

XX PR 09-SEP-1998; 98US-0099643.

XX PR 17-DEC-1998; 98US-0112577.

XX PR 18-DEC-1998; 98US-0112703.

XX PR 22-DEC-1998; 98US-0112933.

XX PR 22-DEC-1998; 98US-0113407.

XX PA (ELIL) LILLY & CO ELI.

XX PI Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;

XX PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;

XX PI Song HY, Wang J, Wu X, Zuckerman SH;

XX DR WPI: 1999-591319/50.

DR N-PSDB; Z25376.

XX Use of mature FLINT for treating acute liver failure, inflammation,

PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic

PT and proinflammatory activity

XX

XX Example 2; Fig 2; 99pp; English.

XX

CC The present invention describes therapeutic applications of mature FLINT

CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT

CC superfamily, which is a member of the tumour necrosis factor receptor

CC family, is used for treating acute liver failure, inflammation of

CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated

CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated

CC injury or disorder such as hypercoagulation (including use with

CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury

CC or disorder, Type I diabetes, cancer, cell damage or damage to an

CC innocent bystander tissue that is induced by a chemotherapeutic agent or

CC therapeutic irradiation, treating haematopoietic progenitor cells that

CC have been exposed to therapeutic radiation or chemotherapy, aplastic

CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is

CC also used for promoting the growth or differentiation of a haematopoietic

CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte

CC resulting from abnormal myocardial ischaemia. The present sequence

CC represents human FLINT.

XX

XX Sequence 302 AA;

DR Query Match 95.6%; Score 131; DB 20; Length 302;

XX Best Local Similarity 96.2%; Pred. No. 3.9e-11;

PT Matches 25; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX

QY 1 PEGWGTPPRAGRAALQLKRLRTEL 26

Db 241 pegwaptpragaaqlklrrrltel 266

XX

RESULT 14

Y29515

ID Y29515 standard; Protein; 527 AA.

XX

XX Y29515;

XX

DT 13-OCT-1999 (first entry)

XX

DE Human lung tumour protein SAL-50 2nd predicted amino acid sequence.

XX

XX Human lung tumour protein; therapy; diagnosis; lung cancer; vaccine;

XX immunotherapy; detection; inhibition.

XX

XX Homo sapiens.

XX

XX WO9938973-A2.

XX

PD 05-AUG-1999.

XX

XX 26-JAN-1999; 99WO-US01642.

XX

XX 22-DEC-1998; 98US-0219245.

XX

XX 28-JAN-1998; 98US-0015022.

XX

XX 28-JAN-1998; 98US-0015029.

XX

XX 18-MAR-1998; 98US-0040828.

XX

XX 18-MAR-1998; 98US-0040831.

XX

XX 23-JUL-1998; 98US-0122191.

XX

XX 23-JUL-1998; 98US-0122192.

XX

XX (CORI-) CORIXA CORP.

XX

XX Frudakis TN, Lodes MJ, Mohamath R, Reed SG;

XX

XX WPI; 1999-479187/40.

DR

DR N-PSDB; Z07194.

XX

XX Lung tumour specific polynucleotides for inhibiting the development

PT of lung cancer

XX

XX Claim 3; Page 170-171; 171pp; English.

XX

XX The present invention describes lung tumour specific polynucleotides

CC and tumour antigens. Z07144 to Z07246 and Z08301 to Z08325 represent

CC specifically claimed polynucleotides, and Y29486 to Y29571 represent

CC amino acid sequences from the present invention. The lung tumour

CC specific polynucleotides and polypeptides can be used in pharmaceutical

CC compositions and vaccines to inhibit the development of lung cancer.

CC They can also be used to detect lung cancer in a patient. Probes and

CC antibodies derived from the lung tumour sequences are useful in

CC detection of lung cancer.

XX

XX Sequence 527 AA;

XX

Query Match 41.6%; Score 57; DB 20; Length 527;

XX Best Local Similarity 47.6%; Pred. No. 2.2;

PT Matches 10; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

XX

QY 1 PEGWGTPPRAGRAALQLKRLR 21

Db 101 peswqptpragaaqlklrrsl 121

XX

RESULT 15

W87502

ID W87502 standard; Protein; 774 AA.

XX

XX W87502;

XX

DT 23-FEB-1999 (first entry)

XX

DE Human N-methyl-D-aspartate receptor subunit encoded by clone NMDA21.

XX

XX Human; N-methyl-D-aspartate receptor; NMDAR2C;

XX NMDA-activated cation-selective ion channel; glutamate receptor.

XX

XX Homo sapiens.

XX

XX U55849895-A.

XX

XX 15-DEC-1998.

XX

XX 20-APR-1994; 94US-0231193.

XX

XX 20-APR-1994; 94US-0231193.

XX

XX 20-APR-1993; 93US-0052449.

XX

XX (SIBI-) SIBIA NEUROSCIENCES INC.

XX

XX Daggett LP, Lu C;

XX

XX WPI; 1999-069812/06.

XX

XX N-PSDB; V82908.

XX

XX DNA encoding N-methyl-D-aspartate receptor subunit - useful for the

XX assembly of functional glutamate receptor subunits

XX

XX Example 3; Columns 249-254; 203pp; English.

XX

XX The present sequence represents a human N-methyl-D-aspartate (NMDA)

XX receptor subunit (NMDAR). The nucleic acid sequence does not contain

XX the 930 5' nucleotides, nor the 51 nucleotides at positions 2351-2401,

XX nor the 1061 3' nucleotides, as set forth in V82889. The cDNA sequence is

XX derived from clone NMDA21. The NMDAR subunits contribute to the formation

XX of NMDA-activated cation-selective ion channels. In addition to being

XX useful for the production of NMDA receptor subunit proteins, the nucleic

XX acids are also useful as probes to identify and isolate nucleic acids

XX encoding related receptor subunits. Functional glutamate receptors can be

CC assembled from several NMDA receptor subunit proteins of one type
CC (homomeric) or from combinations of subunit proteins of different
CC types (heteromeric). The present invention also comprises methods
CC for using such receptor subunits to identify and characterise
CC compounds which affect the function of such receptors, e.g. agonists,
CC antagonists and modulators of glutamate receptor function. The invention
CC also comprises methods for determining whether unknown protein(s) are
CC functional as NMDA receptor subunits.
XX
SQ Sequence 774 AA;

Query Match 39.1%; Score 53.5; DB 20; Length 774;
Best Local Similarity 73.3%; Pred. No. 10;
Matches 11; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

OY 1 PEGWGPTPRAGRAAL 15
| | | | | | | | | |
Db 692 ptgwgpr-pdggraal 705

Search completed: January 30, 2001, 16:45:25
Job time: 585 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:50:27 ; Search time 149.64 seconds
(without alignments)
11.798 Million cell updates/sec

Title: US-09-518-931-2_COPY_239_264

Perfect score: 137

Sequence: 1 PEGWGPTPRAGRAALQLKLRRLTEL 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_66:*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	39.4	126	2 S30707	hypothetical 14.5K
2	51	37.2	112	2 C82664	phage-related prot
3	51	37.2	1249	2 C83358	probable non-ribos
4	50	36.5	680	2 T36472	probable secreted
5	49	35.8	107	2 T49527	hypothetical prote
6	49	35.8	175	1 FRHUL	ferritin light cha
7	49	35.8	668	2 A46013	coagulation factor
8	48.5	35.4	798	2 T19864	hypothetical prote
9	48	35.0	258	2 T42685	hypothetical prote
10	48	35.0	451	2 H75559	conserved hypothet
11	47.5	34.7	322	2 E83009	probable binding p
12	47	34.3	136	2 T08216	outer arm dynein l
13	47	34.3	160	2 A81121	hypothetical prote
14	47	34.3	425	2 JC5909	AE33 protein fru
15	47	34.3	495	2 A75608	aldehyde dehydroge
16	46.5	33.9	232	2 S48353	cdlb protein - she
17	46.5	33.9	333	2 S87246	gene CD1 protein -
18	46.5	33.9	658	2 C82512	hypothetical prote
19	46.5	33.9	1520	2 T00273	hypothetical prote
20	46.5	33.9	2450	2 S71625	protein-tyrosine-p
21	46	33.6	155	2 T25845	hypothetical prote
22	46	33.6	307	2 F71294	hypothetical prote
23	46	33.6	449	1 NBHUHS	complement factor
24	46	33.6	1664	2 H82601	conserved hypothet
25	45.5	33.2	498	2 S11246	LAG-3 protein prec
26	45.5	33.2	1250	1 B45219	N-methyl-D-asparta
27	45	32.8	264	2 T35642	1-acylglycerol-3-p
28	45	32.8	473	2 S09775	hypothetical prote
29	45	32.8	492	2 A61382	phosphorylation re

ALIGNMENTS

RESULT 1

S30707

hypothetical 14.5K protein (corA 3' region) - Escherichia coli

C:Species: Escherichia coli

C:Date: 02-Dec-1993 #sequence_revision 01-Sep-1995 #text_change 04-Mar-2000

C:Accession: S30707; S30744; B65186

R:Danielis, D.L.; Plunkett III, G.; Burland, V.; Blattner, F.R.

Science 257, 771-778, 1992

A:Title: Analysis of the Escherichia coli genome: DNA sequence of the region from 84.

A:Reference number: S30660; MUID:92358234

A:Accession: S30707

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-126 <OHM>

A:Cross-references: EMBL:M87049; NID:9836656; PIDN:AAA67613.1; PID:gl48216

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1992

R:Ohmori, H.

submitted to the EMBL Data Library, January 1993

A:Description: Physical map of the corA region of the E.coli chromosome.

A:Reference number: S30742

A:Accession: S30744

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-126 <OHM>

A:Cross-references: EMBL:L02122; NID:gl45575; PIDN:ADI5039.1; PID:gl45578

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;

.A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617

A:Accession: B55186

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-126 <BLAT>

A:Cross-references: GB:A8000457; GB:U00096; NID:q2367294; PIDN:AAC76820.1; PID:gl7902

A:Experimental source: strain K-12, substrain M61655

C:Genetics:

A:Gene: yigF

A:Map position: 86 min

C:Superfamily: Escherichia coli hypothetical 14.5K protein b3817

Query Match 39.4%; Score 54; DB 2; Length 126;

Best Local Similarity 50.0%; Pred. No. 1.2;

Matches 12; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

Oy 1 PEGWGPTPRAGRAALQLKLRRLT 24

Db 28 PGFWGATPEYKAAFAKVRQLT 51

RESULT 2

C82664
 phase-related protein XF1573 XF1678 [imported] - Xylella fastidiosa (strain 9a5c)
 C:Species: Xylella fastidiosa
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
 C:Accession: C82664; D82650
 R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequence
 Nature 406, 151-157, 2000
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A:Reference number: A82515; MUID:20365717
 A:Note: for a complete list of authors see reference number A59328 below
 A:Accession: C82664
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-112 <SI>
 A:Cross-references: GB:AE003986; GB:AE003849; NID:g9106506; PIDN:AAF84382.1; GSPDB:GN001
 A:Experimental source: strain 9a5c
 A:Accession: D82650
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-112 <SI2>
 A:Cross-references: GB:AE003993; GB:AE003849; NID:g9106734; PIDN:AAF84487.1; GSPDB:GN001
 A:Experimental source: strain 9a5c
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferrelira, A.J.S.
 submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
 Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, F
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
 M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
 A:Reference number: A59328
 A:Contents: annotation
 C:Genetics: <GEN1>
 A:Gene: XF1573
 C:Genetics: <GEN2>
 A:Gene: XF1678

Query Match 37.2%; Score 51; DB 2; Length 112;
 Best Local Similarity 35.0%; Pred. No. 2.9;
 Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 PEGWGTPPRAGRAALQLKLR 20
 Db 41 PDDWKPMPTVGVRGVEIRIR 60

RESULT 3
 C83358
 probable non-ribosomal peptide synthetase PA2305 [imported] - Pseudomonas aeruginosa (st
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 15-Sep-2000
 C:Accession: C83358
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
 ; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
 A:Reference number: A82950
 A:Accession: C83358
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1249 <STO>
 A:Cross-references: GB:AE004656; GB:AE004091; NID:g9948332; PIDN:AAG05693.1; GSPDB:GN001
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA2305

Query Match 37.2%; Score 51; DB 2; Length 1249;
 Best Local Similarity 57.1%; Pred. No. 33;
 Matches 12; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 PEGWGTPPRAGRAALQLKLR 21
 Db 810 PEGDEPEPSAGAAGLSAERR 830

RESULT 4
 T36472
 probable secreted alpha-galactosidase - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T36472
 R:Seeger, K.J.; Harris, D.; Thomson, N.R.; Parkhill, J.; Barrell, B.G.; Rajandream, M
 submitted to the EMBL Data Library, September 1999
 A:Reference number: Z1607
 A:Accession: T36472
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-680 <SEE>
 A:Cross-references: EMBL:AL110470; PIDN:CAE54169.1; GSPDB:GN00070; SCOEDB:SCF85.12
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOEDB:SCF85.12

Query Match 36.5%; Score 50; DB 2; Length 680;
 Best Local Similarity 50.0%; Pred. No. 25;
 Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 PEGWGTPPRAGRAALQLK 18
 Db 464 PDGWRVTPRSATSARALK 481

RESULT 5
 T49527
 hypothetical protein B21J21.60 [imported] - Neurospora crassa
 C:Species: Neurospora crassa
 C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
 C:Accession: T49527
 R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatu
 submitted to the Protein Sequence Database, May 2000
 A:Reference number: Z25022
 A:Accession: T49527
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-107 <SCH>
 A:Cross-references: EMBL:AL355929; GSPDB:GN00116; NCSP:B21J21.60
 A:Experimental source: BAC clone B21J21; strain OR74A
 C:Genetics:
 A:Gene: NCSP:B21J21.60
 A:Map position: 6

Query Match 35.8%; Score 49; DB 2; Length 107;
 Best Local Similarity 47.6%; Pred. No. 5.3;
 Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 PEGWGTPPRAGRAALQLKLR 21
 Db 40 PWCWGPTWDARSSPQMER 60

RESULT 6
 FRHUL
 ferritin light chain - human
 C:Species: Homo sapiens (man)
 C:Date: 13-Jun-1983 #sequence_revision 03-May-1996 #text_change 03-Mar-2000
 C:Accession: B23920; B24844; A91308; A91308; I37137; S68943; A03266
 R:Boyd, D.; Vecoli, C.; Belcher, D.M.; Jain, S.K.; Drysdale, J.W.

QY 3 GWGPTPRAGRA 13

A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.

A:Reference number: A81000; MUID:20175755

A:Accession: A81121

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-160 <TET>

A:Cross-references: GB:AE002460; GB:AE002098; NID:g7226335; PIDN:AAF41491.1; PID:g722633

A:Experimental source: serogroup B, strain MMD58

C:Genetics:

A:Gene: NWB1100

Query Match

Best Local Similarity 34.3%; Score 47; DB 2; Length 160;

Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 11 GRAALQLKRRRLTE 25

I: |||::|||:

Db 11 GRVALQMRIRRLKQ 25

RESULT 14

JC5909

AE33 protein - fruit fly (*Drosophila melanogaster*)

C:Species: *Drosophila melanogaster*

C>Date: 18-Mar-1998 #sequence_revision 18-Mar-1998 #text_change 07-May-1999

C:Accession: JC5909

R:Demille, M.M.C.; Kimmel, B.E.; Rubin, G.M.

Gene 183, 103-108, 1996

A:Title: A *Drosophila* gene regulated by rough and glass shows similarity to ena and VASH

A:Reference number: JC5909; MUID:97149285

A:Accession: JC5909

A:Molecule type: mRNA

A:Residues: 1-425 <DEM>

A:Comment: This protein plays a role in signalling during R2 and R5 development.

C:Genetics:

A:Gene: ro

C:Keywords: phosphoprotein

F:262/Binding site: phosphate (Ser) (covalent) (by cAMP- and cGMP-dependent kinases) #st

F:311/Binding site: phosphate (Thr) (covalent) (by cAMP- and cGMP-dependent kinases) #st

Query Match

Best Local Similarity 34.3%; Score 47; DB 2; Length 425;

Matches 10; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 2 EGWGPPTPRAGRAALQLKRRRLTEL 26

||| | | | | | | | | |

Db 25 EGWLPLAGGGLANVSIRKRLSPL 49

RESULT 15

A75608

aldehyde dehydrogenase - *Deinococcus radiodurans* (strain R1)

C:Species: *Deinococcus radiodurans*

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C:Accession: A75608

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.

A:Reference number: A75250; MUID:20036896

A:Accession: A75608

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-495 <WHI>

A:Cross-references: GB:AE001862; GB:AE001825; NID:g6460468; PIDN:AAF12231.1; PID:g646052

A:Experimental source: strain R1

C:Genetics:

A:Gene: DRA0126

A:Map position: 2

C:Superfamily: aldehyde dehydrogenase (NAD+); aldehyde dehydrogenase homology

Query Match

Best Local Similarity 34.3%; Score 47; DB 2; Length 495;

Matches 10; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 2 EGWGPPTPRAGRAALQLKRRRLTE 25

: | | | | | | | | | |

Db 74 DSWSQTPAEERAAALLERISRGLE 97

Search completed: January 30, 2001, 16:50:29

Job time: 719 sec

us-09-518-931-2_copy_239_264.rpr

Wed Jan 31 13:44:15 2001

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	54	39.4	126	1	YIGF_ECOLI	P27842	escherichia
2	53.5	39.1	1233	1	NME3_HUMAN	Q14957	homo sapien
3	49	35.8	174	1	FRIL_HUMAN	P02792	homo sapien
4	49	35.8	668	1	F13B_MOUSE	Q07968	mus musculus
5	46	33.6	307	1	V678_TREPA	O83684	treponema p
6	46	33.6	3321	1	KEND_HUMAN	O95613	homo sapien
7	45.5	33.2	525	1	LAC3_HUMAN	P18627	homo sapien
8	45.5	33.2	1237	1	NME3_RAT	Q00961	rattus norv
9	45	32.8	473	1	UL13_HCMVA	P16755	human cytom
10	45	32.8	676	1	ICP0_HSVBK	P29836	bovine herp
11	45	32.8	928	1	VGLB_HSVBP	P17471	bovine herp
12	45	32.8	1221	1	FBL2_MOUSE	P37889	mus musculus
13	44.5	32.5	450	1	AGOA_MYCTU	P22487	mycobacteri
14	44.5	32.5	537	1	GAG_MLVRD	P11269	radiation m
15	44.5	32.5	3176	1	CA36_HUMAN	P12111	homo sapien
16	44	32.1	174	1	FRIL_HORSE	P2791	equus cabal
17	44	32.1	244	1	F1MB_BORPE	P33409	bordetella
18	44	32.1	270	1	CFHE_HUMAN	P36980	homo sapien
19	44	32.1	330	1	CPH1_HUMAN	Q03591	homo sapien
20	44	32.1	374	1	PAB_MYCTU	P15712	mycobacteri
21	44	32.1	418	1	GSA_ARCFO	O29027	archaeoglob
22	44	32.1	715	1	AT12_HSVIF	P08314	herpes simp
23	44	32.1	1286	1	CTRO_HUMAN	O14578	homo sapien
24	44	32.1	1409	1	RPOC_PSEPU	P19176	pseudomonas
25	44	32.1	1597	1	CTRO_MOUSE	P49025	mus musculus
26	44	32.1	4385	1	YP73_CAEEL	Q09222	caenorhabdi
27	43.5	31.8	812	1	EBM3_EBV	P12977	epstein-bar
28	43	31.4	162	1	CRF1_CATCO	P13241	catostomus
29	43	31.4	376	1	DCOP_CEPAC	P14017	cephalospor
30	43	31.4	516	1	YC97_BRAJA	Q45221	bradyrhizob
31	43	31.4	676	1	ICP0_HSVBJ	P29128	bovine herp
32	43	31.4	780	1	PRTP_HSVIF	P06490	herpes simp
33	43	31.4	785	1	PRTP_HSV11	P10212	herpes simp

DR PFAM; PF00210; ferritin; 1.
DR PROSITE; PS00204; FERRITIN_2; 1.
DR PROSITE; PS00540; FERRITIN_1; 1.
KW Iron storage; Multigene family; Acetylation.
FT INIT_MET 0 0
FT MOD_RES 1 1 ACETYLATION.
FT DOMAIN 53 60 CATALYTIC SITE FOR IRON OXIDATION.
FT METAL 53 53 IRON (POTENTIAL).
FT METAL 56 56 IRON (POTENTIAL).
FT METAL 57 57 IRON (POTENTIAL).
FT METAL 60 60 IRON (POTENTIAL).
FT METAL 63 63 IRON (POTENTIAL).
FT CONFLICT 53 53 E -> Q (IN REF. 4).
FT CONFLICT 86 86 E -> Q (IN REF. 4).
FT CONFLICT 101 101 A -> T (IN REF. 2).
FT CONFLICT 174 174 D -> N (IN REF. 4).
SQ SEQUENCE 174 AA; 19888 MW; 8F0B4B23B6CAFEF2 CRC64;

Query Match 35.8%; Score 49; DB 1; Length 174;
Best Local Similarity 48.0%; Pred. No. 2.6;
Matches 12; Conservative 4; Mismatches 7; Indels 2; Gaps 1;

Oy 4 WGPTRAGRAALQL--KLRRRLTEL 26
||| | :|| :| | :| |
Db 89 WCKTPDAMKAAMALEKLNQALLDL 113

RESULT 4
F13B_MOUSE STANDARD; PRT; 668 AA.
AC Q07968;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE COAGULATION FACTOR XIII B CHAIN PRECURSOR (EC 2.3.2.13) (PROTEIN-
DE GLUTAMINE GAMMA-GLUTAMYLTRANSFERASE B CHAIN) (TRANSGLUTAMINASE B
DE CHAIN).
GN F13B OR CF13B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=B10.D2/OSN; TISSUE=LIVER;
RX MEDLINE=93224141; PubMed=8468048;
RA Nonaka M., Matsuda Y., Shirosaki T., Moriwaki K., Nonaka M.,
RA Natsume-Sakai S.;
RT "Molecular cloning of the b subunit of mouse coagulation factor XIII
RT and assignment of the gene to chromosome 1: close evolutionary
RL relationship to complement factor H";
RL Genomics 15:535-542(1993).
CC -!- FUNCTION: THE B CHAIN OF FACTOR XIII IS NOT CATALYTICALLY ACTIVE,
CC BUT IS THOUGHT TO STABILIZE THE A SUBUNITS AND REGULATE THE RATE
CC OF TRANSGLUTAMINASE FORMATION BY THROMBIN.
CC -!- SUBUNIT: TETRAMER OF TWO A CHAINS AND TWO B CHAINS.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN LIVER AND KIDNEY.
CC -!- SIMILARITY: CONTAINS 10 SUSHI (SCR) REPEATS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; D10071; BAA00963.1; -.
DR PIR; A46013; A46013.
DR HSP; P10998; 1VVC.
DR MGD; MGI:88379; F13B.
DR INTERPRO; IPR000436; -.
DR PFAM; PF00084; sushi; 8.

KW Transferase; Plasma; Blood coagulation; Repeat; Glycoprotein; Signal;
KW Sushi.
FT SIGNAL 1 20 BY SIMILARITY.
FT CHAIN 21 668 COAGULATION FACTOR XIII B CHAIN.
FT DOMAIN 24 647 10 X SUSHI (SCR) REPEATS.
FT REPEAT 24 88 SUSHI 1.
FT REPEAT 90 147 SUSHI 2.
FT REPEAT 152 209 SUSHI 3.
FT REPEAT 212 268 SUSHI 4.
FT REPEAT 273 328 SUSHI 5.
FT REPEAT 335 390 SUSHI 6.
FT REPEAT 395 451 SUSHI 7.
FT REPEAT 453 516 SUSHI 8.
FT REPEAT 523 579 SUSHI 9.
FT REPEAT 581 647 SUSHI 10.
FT DISULFID 25 76 BY SIMILARITY.
FT DISULFID 59 87 BY SIMILARITY.
FT DISULFID 91 135 BY SIMILARITY.
FT DISULFID 118 146 BY SIMILARITY.
FT DISULFID 153 197 BY SIMILARITY.
FT DISULFID 180 208 BY SIMILARITY.
FT DISULFID 213 255 BY SIMILARITY.
FT DISULFID 241 267 BY SIMILARITY.
FT DISULFID 274 316 BY SIMILARITY.
FT DISULFID 302 327 BY SIMILARITY.
FT DISULFID 336 378 BY SIMILARITY.
FT DISULFID 364 389 BY SIMILARITY.
FT DISULFID 396 439 BY SIMILARITY.
FT DISULFID 425 450 BY SIMILARITY.
FT DISULFID 454 505 BY SIMILARITY.
FT DISULFID 486 515 BY SIMILARITY.
FT DISULFID 524 567 BY SIMILARITY.
FT DISULFID 553 578 BY SIMILARITY.
FT DISULFID 582 636 BY SIMILARITY.
FT DISULFID 616 646 BY SIMILARITY.
FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 545 545 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 668 AA; 76078 MW; 80BC9E00A9E53FA6 CRC64;

Query Match 35.8%; Score 49; DB 1; Length 668;
Best Local Similarity 47.4%; Pred. NO. 11;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Oy 2 ECGGTPPRAGRAALQLKLR 20
||| | | :| |
Db 79 ECGSPNRCYKCKLPDLR 97

RESULT 5
Y678_TREPA STANDARD; PRT; 307 AA.
ID Y678_TREPA
AC O83684;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYPOTHETICAL PROTEIN TP0678.
GN TP0678.
OS Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
RC Bacteria; Spirochaetales; Treponema.
RC SEQUENCE FROM N.A.
RC STRAIN=NICHOLS;
RX MEDLINE=98332770; PubMed=9665876;
RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uitterback T.,
RA McDonald L., Artlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.";

[illegible]

CC between the Swiss Institute of Bioinformatics. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X51985; CAA36243.1; -;
DR EMBL; A21333; CAA01547.1; ALT_SEQ.
DR PIR; S11246; S11246.
DR MIM; 153337; -.


```

FT CARBOHYD 356 356 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 473 AA; 54613 MW; 378C32A16D3F018 CRC64;

Query Match 32.8%; Score 45; DB 1; Length 473;
Best Local Similarity 48.1%; Pred. No. 29;
Matches 13; Conservative 4; Mismatches 6; Indels 4; Gaps 2;

QY 1 PEGGPT---PRAGRAALQLKRRRLT 24
| | | | | | | | | | | | | | | | | | | | | |
Db 238 PAGSGPNTGGPRPGAGLR-QLRQQLT 263

RESULT 10
ICPO_HSVBK STANDARD; PRT; 676 AA.
ID ICPO_HSVBK
AC P29836;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE TRANS-ACTING TRANSCRIPTIONAL PROTEIN ICPO (P135 PROTEIN) (IER
DE 2.9/BR2.6).
GN BICPO.
OS Bovine herpesvirus type 1 (strain K22).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92219360; PubMed=1313901;
RA Wirth U.V., Fraefel C., Vogt B., Vicek C., Paces V., Schwyzler M.;
RT "Immediate-early RNA 2.9 and early RNA 2.6 of bovine herpesvirus 1
RT are 3' coterminal and encode a putative zinc finger transactivator
RT protein.";
RL J. Virol. 66:2763-2772(1992).
CC -|- PTM: THE STRONGLY ACIDIC REGION MIGHT SERVE AS A TRANSCRIPTIONAL
CC ACTIVATION DOMAIN, POSSIBLY REGULATED THROUGH PHOSPHORYLATION BY
CC CASEIN KINASE II.
CC -|- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.
CC -|- SIMILARITY: TO OTHER HERPESVIRUSES ICPO PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M84464; AAA46061.1; -.
CC PIR; A38209; EDBE22.
CC DR HSPP; P28990; ICHC.
CC DR INTERPRO: IPR001841; -.
CC DR PFAM: PF00097; zf-C3HC4; 1.
CC DR PROSITE; PS00518; ZINC_FINGER_C3HC4; 1.
KW Transcription regulation; Trans-acting factor; Activator; Zinc-finger;
KW DNA-binding; Early protein; Repressor; Phosphorylation.
FT ZN_FING 13 51
FT DOMAIN 284 331 ASP/GLU-RICH (ACIDIC).
FT SEQUENCE 676 AA; 67701 MW; 9BB0683C9BFC65D CRC64;

Query Match 32.8%; Score 45; DB 1; Length 676;
Best Local Similarity 36.1%; Pred. No. 42;
Matches 13; Conservative 1; Mismatches 10; Indels 12; Gaps 1;

QY 3 GWTPTPRAGRAALQLKL-----RRRLTEL 26
| | | | | | | | | | | | | | | | | | | | | |
Db 434 GWTGSEGEGRRAKLGLGEAGPRVQARRRTEL 469

RESULT 11
VGLB_HSVBP STANDARD; PRT; 928 AA.
ID VGLB_HSVBP

```

```

AC P17471;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE GLYCOPROTEIN I PRECURSOR (GLYCOPROTEIN GVP-6) (GLYCOPROTEIN 11A)
DE (GLYCOPROTEIN 16) (GLYCOPROTEIN G130).
GN Gi.
OS Bovine herpesvirus type 1 (strain P8-2).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89020821; PubMed=28456660;
RA Misra V., Nelson R., Smith M.;
RT "Sequence of a bovine herpesvirus type-1 glycoprotein gene that is
RT homologous to the herpes simplex gene for the glycoprotein gB.";
RL Virology 166:542-549(1988).
CC -|- SUBUNIT: DIMER, PROBABLY LINKED BY DISULFIDE BONDS.
CC -|- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M23257; AAA46013.1; -.
CC PIR; A31166; VGBEBG.
CC DR INTERPRO: IPR000234; -.
CC DR PFAM: PF00606; Glycoprotein_B; 2.
KW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 67
FT CHAIN 68 928 GLYCOPROTEIN I.
FT DOMAIN 68 780 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 781 801 POTENTIAL.
FT DOMAIN 804 824 POTENTIAL.
FT CARBOHYD 105 105 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 153 153 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 484 484 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 579 579 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 703 703 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 409 409 S -> T (IN AAA46013).
FT CONFLICT 673 673 G -> P (IN AAA46013).
SQ SEQUENCE 928 AA; 102177 MW; B47982224FCD769D CRC64;

Query Match 32.8%; Score 45; DB 1; Length 928;
Best Local Similarity 40.6%; Pred. No. 58;
Matches 13; Conservative 2; Mismatches 7; Indels 10; Gaps 2;

QY 4 WGP-----TPRAGRAAL-----QLKLRRLTE 25
| | | | | | | | | | | | | | | | | | | | | |
Db 597 WGTGHRDVLPRAGRGALFIENSARARRLLQ 628

RESULT 12
FBL2_MOUSE STANDARD; PRT; 1221 AA.
ID FBL2_MOUSE
AC P37889; Q9WUI2;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE FIBULIN-2 PRECURSOR.
GN FBLN2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]

```

RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-35.
RC TISSUE=FIBROBLAST;
RX MEDLINE=94064787; PubMed=9245130;
RA Pan T.-C., Sasaki T., Zhang R.-Z., Faessler R., Timpl R., Chu M.-L.;
RT "Structure and expression of fibulin-2, a novel extracellular matrix
protein with multiple EGF-like repeats and consensus motifs for
calcium binding.";
RL J. Cell Biol. 123:1269-1277(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99337686; PubMed=10406956;
RA Gressel S., Sicot F.-X., Gotta S., Chu M.-L.;
RT "Mouse fibulin-2 gene. Complete exon-intron organization and promoter
characterization.";
RL Eur. J. Biochem. 263:471-477(1999).
RN [1]
CC -1- FUNCTION: ITS BINDING TO FIBRONECTIN AND SOME OTHER LIGANDS IS
CALCIUM DEPENDENT.
CC -1- SUBUNIT: DISULFIDE-LINKED HOMOTRIMER.
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST TWO FORMS OF FIBULIN-2 DUE TO THE
PRESENCE OR ABSENCE OF A SINGLE EGF-LIKE (3) REPEAT MAY ARISE BY
ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: COMPONENT OF BOTH BASEMENT MEMBRANES AND OTHER
CONNECTIVE TISSUES.
CC -1- SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 11 EGF-LIKE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X75285; CAA53040.1; -;
DR EMBL; AF135253; AAD34456.1; -;
DR EMBL; AF135239; AAD34456.1; JOINED.
DR EMBL; AF135240; AAD34456.1; JOINED.
DR EMBL; AF135241; AAD34456.1; JOINED.
DR EMBL; AF135242; AAD34456.1; JOINED.
DR EMBL; AF135243; AAD34456.1; JOINED.
DR EMBL; AF135244; AAD34456.1; JOINED.
DR EMBL; AF135245; AAD34456.1; JOINED.
DR EMBL; AF135246; AAD34456.1; JOINED.
DR EMBL; AF135247; AAD34456.1; JOINED.
DR EMBL; AF135248; AAD34456.1; JOINED.
DR EMBL; AF135249; AAD34456.1; JOINED.
DR EMBL; AF135250; AAD34456.1; JOINED.
DR EMBL; AF135251; AAD34456.1; JOINED.
DR EMBL; AF135252; AAD34456.1; JOINED.
DR PIR; A49457; A49457.
DR HSP; P07204; IFGD.
DR MGD; MG1:95488; FBLN2.
DR INTERPRO; IPR000020; -;
DR INTERPRO; IPR000152; -;
DR INTERPRO; IPR000561; -;
DR INTERPRO; IPR001881; -;
DR PFAM; PF01821; ANATO; 2.
DR PFAM; PF00008; EGF; 6.
DR PROSITE; PS00010; ASX_HYDROXYL; 5.
DR PROSITE; PS01177; ANAPHYLATOXIN_1; 3.
DR PROSITE; PS01178; ANAPHYLATOXIN_2; 3.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 5.
DR PROSITE; PS01187; EGF_CA; 10.
KW Signal; Glycoprotein; Extracellular matrix; Plasma; EGF-like domain;
Calcium-binding; Alternative splicing; Repeat.
FT SIGNAL 1 26
FT CHAIN 27 1221 FIBULIN-2.
FT DOMAIN 27 434 N.
FT DOMAIN 27 176 SUBDOMAIN NA (CYS-RICH).
FT DOMAIN 177 434 SUBDOMAIN NB (CYS-FREE).
FT DOMAIN

FT DOMAIN 435 543 3 X ANAPHYLATOXIN REPEATS.
FT REPEAT 435 477 ANAPHYLATOXIN-LIKE 1.
FT REPEAT 478 510 ANAPHYLATOXIN-LIKE 2.
FT REPEAT 511 543 ANAPHYLATOXIN-LIKE 3.
FT DOMAIN 594 1106 11 X EGF-TYPE REPEATS (DOMAIN II).
FT DOMAIN 594 635 EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 669 708 EGF-LIKE 2.
FT DOMAIN 709 755 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 756 800 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 801 846 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 847 894 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 895 937 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 938 979 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 980 1018 EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1019 1061 EGF-LIKE 10, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1062 1106 EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1111 1221 DOMAIN III.
FT SITE 421 423 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 435 462 BY SIMILARITY.
FT DISULFID 436 469 BY SIMILARITY.
FT DISULFID 449 470 BY SIMILARITY.
FT DISULFID 479 508 BY SIMILARITY.
FT DISULFID 492 509 BY SIMILARITY.
FT DISULFID 511 535 BY SIMILARITY.
FT DISULFID 512 542 BY SIMILARITY.
FT DISULFID 525 543 BY SIMILARITY.
FT DISULFID 598 610 BY SIMILARITY.
FT DISULFID 606 619 BY SIMILARITY.
FT DISULFID 621 634 BY SIMILARITY.
FT DISULFID 673 683 BY SIMILARITY.
FT DISULFID 679 692 BY SIMILARITY.
FT DISULFID 694 707 BY SIMILARITY.
FT DISULFID 713 726 BY SIMILARITY.
FT DISULFID 720 735 BY SIMILARITY.
FT DISULFID 742 754 BY SIMILARITY.
FT DISULFID 805 818 BY SIMILARITY.
FT DISULFID 812 827 BY SIMILARITY.
FT DISULFID 833 845 BY SIMILARITY.
FT DISULFID 899 912 BY SIMILARITY.
FT DISULFID 906 921 BY SIMILARITY.
FT DISULFID 923 936 BY SIMILARITY.
FT DISULFID 942 954 BY SIMILARITY.
FT DISULFID 950 963 BY SIMILARITY.
FT DISULFID 965 978 BY SIMILARITY.
FT DISULFID 984 993 BY SIMILARITY.
FT DISULFID 989 1002 BY SIMILARITY.
FT DISULFID 1004 1017 BY SIMILARITY.
FT DISULFID 1023 1035 BY SIMILARITY.
FT DISULFID 1031 1044 BY SIMILARITY.
FT DISULFID 1046 1060 BY SIMILARITY.
FT DISULFID 1066 1079 BY SIMILARITY.
FT DISULFID 1073 1088 BY SIMILARITY.
FT DISULFID 1093 1105 BY SIMILARITY.
FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 497 497 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 737 737 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1072 1072 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 709 755 MISSING (IN ISOFORM EGF3-LESS).
FT CONFLICT 140 159 HSGRKYAGHTVHLSSCRAC -> TVAVSICWPYRPLILP
GF (IN REF. 2).
FT CONFLICT 348 348 S -> L (IN REF. 2).
FT CONFLICT 507 507 Q -> QQ (IN REF. 2).
FT CONFLICT 1102 1102 Q -> E (IN REF. 2).
SQ SEQUENCE 1221 AA; 131818 MW; 87DB2A10A8FDC45F CRC64;

Query Match 32.8%; Score 45; DB 1; Length 1221;
Best Local Similarity 55.6%; Pred. No. 78;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
Qy 2 EGVGTPPRAGRAALQLKL 19
||| ||| ||| |
Db 340 EGAAVPRSGLAALSPSL 357

Db 427 QGWGYQPRSGQRA-----RRAT 444

RESULT 14

ID GAG_MLVRO STANDARD; PRT; 537 AA.

AC P11269;

DT 01-JUL-1989 (Rel. 11, Created)

DT 01-JUL-1989 (Rel. 11, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE GAG POLYPROTEIN [CONTAINS: CORE PROTEIN P15; INNER COAT PROTEIN P12;

DE CORE SHELL PROTEIN P30; NUCLEOPROTEIN P10].

GN GAG.

OS Radiation murine leukemia virus.

OC Viruses; Retroviral viruses; Retroviridae; Mammalian type C retroviruses.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=87207680; PubMed=3033897;

RA Merregaert J., Janowski M., Reddy E.P.;

RT "Nucleotide sequence of a radiation leukemia virus genome.";

RL Virology 158:88-102(1987).

RN [2]

RP SEQUENCE OF 478-505.

RX MEDLINE=81264245; PubMed=6267042;

RA Henderson L.E., Copeland T.D., Sowder R.C., Smythers G.W.,

RT Oroszlan S.;

RT "Primary structure of the low molecular weight nucleic acid-binding

RL proteins of murine leukemia viruses.";

RL J. Biol. Chem. 256:8400-8406(1981).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; K03363; AAA46518.1;

DR PIR; A26183; FOMVRV.

DR INTERPRO; IPR000840;

DR INTERPRO; IPR002079;

DR INTERPRO; IPR003036;

DR PFAM; PF02093; Gag_p30; 1.

DR PFAM; PF01140; gag_MA; 1.

DR PFAM; PF01141; gag_p12; 1.

DR Core protein; Coat protein; Nucleoprotein; Polyprotein; Myristate.

KW CHAIN 2 129

FT CHAIN 130 214

FT CHAIN 215 477

FT CHAIN 478 537

FT LIPID 2 2

FT MYRISTATE.

FT CONFLICT 479 479 T -> S (IN REF. 2).

FT CONFLICT 482 482 T -> S (IN REF. 2).

FT SEQUENCE 537 AA; 60784 MW; 312AF7B2BBB4B7FB CRC64;

Query Match 32.5%; Score 44.5; DB 1; Length 537;

Best Local Similarity 41.7%; Pred. No. 39;

Matches 10; Conservative 5; Mismatches 8; Indels 1; Gaps 1;

QY 2 EGVGPTPR-AGRAALQQLKRRRLT 24

DB 14 EHWGDVQRIASNSQSVKRRRV 37

RESULT 15

CA36_HUMAN

ID CA36_HUMAN STANDARD; PRT; 3176 AA.

AC P12111; Q16501;

DT 01-OCT-1989 (Rel. 12, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

RESULT 13

AROA_MYCTU STANDARD; PRT; 450 AA.

AC P22487;

DT 01-AUG-1991 (Rel. 19, Created)

DT 01-AUG-1991 (Rel. 19, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE (EC 2.5.1.19) (5-

DE ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE SYNTHASE) (EPSP SYNTHASE) (EPSPS).

GN AROA OR RV3227 OR MTCV20B11.02.

OS Mycobacterium tuberculosis.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=H37RV;

RX MEDLINE=91072223; PubMed=2123856;

RA Garbe T., Jones C., Charles I.G., Dougan G., Young D.;

RT "Cloning and characterization of the aroA gene from Mycobacterium

RL tuberculosis.";

RL J. Bacteriol. 172:6774-6782(1990).

RN [2]

RP SEQUENCE FROM N.A.

RX STRAIN=H37RV;

RX MEDLINE=98295987; PubMed=9634230;

RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,

RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,

RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,

RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,

RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,

RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,

RA Rutter K., Seeger K., Skelton S., Squares S., Sultston J.E.,

RA Taylor K., Whitehead S., Barrrell B.G.;

RT "Deciphering the biology of Mycobacterium tuberculosis from the

RT complete genome sequence.";

RL Nature 393:537-544(1998).

CC -1- CATALYTIC ACTIVITY: PHOSPHOENOLPYRUVATE + 3-PHOSPHOSHIKIMATE =

CC ORTHOPHOSPHATE + O(5)-(1-CARBOXYVINYL)-3-PHOSPHOSHIKIMATE.

CC -1- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN

CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).

CC -1- SUBUNIT: MONOMER.

CC -1- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; X52269; CA36510.1;

DR EMBL; M62708; AAA23556.1;

DR EMBL; Z95121; CAB08328.1;

DR PIR; A37807; A37807.

DR TUBERCULIST; RV3227;

DR INTERPRO; IPR001986;

DR PFAM; PF00275; EPSP_synthase; 1.

DR PROSITE; PS00104; EPSP_SYNTHASE_1; 1.

DR PROSITE; PS00885; EPSP_SYNTHASE_2; 1.

KW Aromatic amino acid biosynthesis; Transferase.

QY SEQUENCE 450 AA; 46425 MW; 27BB86F9412A07D5 CRC64;

Query Match 32.5%; Score 44.5; DB 1; Length 450;

Best Local Similarity 45.8%; Pred. No. 32;

Matches 11; Conservative 3; Mismatches 3; Indels 7; Gaps 2;

QY 2 EGVG-PTPRAGRAALQQLKRRRLT 24

:||| | |||: | ||

DE COLLAGEN ALPHA 3(VI) CHAIN PRECURSOR.
GN COL6A3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=FIBROBLAST;
RX MEDLINE=90151612; PubMed=1689238;
RA Chu M.-L., Zhang R.-Z., Pan T.-C., Stokes D., Conway D., Kuo H.-J.,
RA Glanville R., Mayer U., Mann K., Deutzmann R., Timpl R.;
RT "Mosaic structure of globular domains in the human type VI collagen
RT alpha 3 chain: similarity to von Willebrand factor, fibronectin,
RT actin, salivary proteins and aprotinin type protease inhibitors.";
RL EMBO J. 9:385-393(1990).
RN [2]
RP REVISIONS.
RA Chu M.-L.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 2038-2373 FROM N.A.
RX MEDLINE=89066644; PubMed=3198591;
RA Chu M.-L., Conway D., Pan T.-C., Baldwin C., Mann K., Deutzmann R.,
RA Timpl R.;
RT "Amino acid sequence of the triple-helical domain of human collagen
RT type VI.";
RL J. Biol. Chem. 263:18601-18606(1988).
RN [4]
RP SEQUENCE OF 2092-2157 FROM N.A.
RX MEDLINE=88029444; PubMed=3665927;
RA Chu M.-L., Mann K., Deutzmann R., Pribula-Conway D.,
RA Hsu-Chen C.-C., Bernard M.P., Timpl R.;
RT "Characterization of three constituent chains of collagen type VI by
RT peptide sequences and cDNA clones.";
RL Eur. J. Biochem. 168:309-317(1987).
RN [5]
RP SEQUENCE OF 2092-2151 FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=88161046; PubMed=3348212;
RA Weil D., Mattei M.-G., Passage E., van Cong N., Pribula-Conway D.,
RA Mann K., Deutzmann R., Timpl R., Chu M.-L.;
RT "Cloning and chromosomal localization of human genes encoding the
RT three chains of type VI collagen.";
RL Am. J. Hum. Genet. 42:435-445(1988).
RN [6]
RP SEQUENCE OF 32-236 FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=93054780; PubMed=1339440;
RA Zanussi S., Dollana R., Segat D., Bonaldo P., Colombatti A.;
RT "The human type VI collagen gene. mRNA and protein variants of the
RT alpha 3 chain generated by alternative splicing of an additional 5-end
RT exon.";
RL J. Biol. Chem. 267:24082-24089(1992).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS) OF 3107-3164.
RX MEDLINE=95182468; PubMed=7533217;
RA Arnoux B., Merigau K., Saludjian P., Norris F., Norris K., Bjoern S.,
RA Olsen O., Petersen L., Ducruix A.;
RT "The 1.6 A structure of Kunitz-type domain from the alpha 3 chain of
RT human type VI collagen.";
RL J. Mol. Biol. 246:609-617(1995).
RN [8]
RP STRUCTURE BY NMR OF 3102-3164.
RX MEDLINE=96398604; PubMed=8805527;
RA Zwickstetter M., Czisch M., Mayer U., Chu M.-L., Zinth W., Timpl R.,
RA Holak T.A.;
RT "Structure and multiple conformations of the kunitz-type domain from
RT human type VI collagen alpha3(VI) chain in solution.";
RN [9]
RP STRUCTURE BY NMR OF 3107-3164.
RX MEDLINE=97410331; PubMed=9265624;
RA Soerensen M.D., Bjoern S., Norris K., Olsen O., Petersen L.,
RA James T.L., Led J.J.;

"Solution structure and backbone dynamics of the human alpha3-chain
type VI collagen C-terminal Kunitz domain.";
Biochemistry 36:10439-10450(1997).
[10]
RP VARIANT BM GLU-1679, AND VARIANT HIS-2831.
RX MEDLINE=98204804; PubMed=9536084;
RA Pan T.-C., Zhang R.-Z., Pericak-Vance M.A., Tandan R., Fries T.,
RA Stajich J.M., Viles K., Vance J.M., Chu M.-L., Speer M.C.;
RT "Missense mutation in a von Willebrand factor type A domain of the
RT alpha 3(VI) collagen gene (COL6A3) in a family with Bethlem
RT myopathy.";
RL Hum. Mol. Genet. 7:807-812(1998).
CC -!- FUNCTION: COLLAGEN VI ACTS AS A CELL-BINDING PROTEIN.
CC -!- SUBUNIT: TRIMERS COMPOSED OF THREE DIFFERENT CHAINS: ALPHA 1(VI),
CC ALPHA 2(VI), AND ALPHA 3(VI).
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -!- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC -!- DISEASE: DEFECTS IN COL6A3 ARE A CAUSE OF BETHLEM MYOPATHY (BM).
CC BM IS A RARE AUTOSOMAL DOMINANT PROXIMAL MYOPATHY CHARACTERIZED BY
CC EARLY CHILDHOOD ONSET (COMPLETE PENETRATION BY THE AGE OF 5) AND
CC JOINT CONTRACTURES MOST FREQUENTLY AFFECTING THE ELBOWS AND
CC ANKLES.
CC -!- SIMILARITY: CONTAINS 1 PROTEASE INHIBITOR DOMAIN BELONGING TO THE
CC BPTI/KUNITZ FAMILY OF INHIBITORS
CC -!- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
CC or send an email to license@isb-sib.ch.

DR EMBL; X52022; CAA36267.1; -.
DR EMBL; X06196; CAA29557.1; -.
DR EMBL; M20778; -; NOT_ANNOTATED_CDS.
DR EMBL; M27449; AAS52057.1; -.
DR EMBL; S49432; AAB24261.1; -.
DR PIR; C31952; C31952.
DR PDB; 1KNT; 01-NOV-94.
DR PDB; 2KNT; 15-MAY-97.
DR PDB; 1KUN; 12-NOV-97.
DR MIM; 120250; -.
DR MIM; 158810; -.
DR INTERPRO; IPR0000087; -.
DR INTERPRO; IPR001777; -.
DR INTERPRO; IPR002035; -.
DR INTERPRO; IPR002223; -.
DR PFAM; PF01391; Collagen; 5.
DR PFAM; PF00014; Kunitz_BPTI; 1.
DR PFAM; PF00041; fn3; 1.
DR PFAM; PF00092; vwa; 11.
DR PRINTS; PR00453; VWFADOMAIN.
DR PRINTS; PR00759; BASICPTASE.
DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR PROSITE; PS02279; BPTI_KUNITZ_2; 1.
KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Glycoprotein; Cell adhesion; Serine protease inhibitor; Collagen;
KW Signal; 3D-structure; Disease mutation; Polymorphism;
KW Alternative splicing.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 3176 COLLAGEN ALPHA 3(VI) CHAIN.
FT DOMAIN 26 2038 NONHELICAL REGION.
FT DOMAIN 2039 2375 TRIPLE-HELICAL REGION.
FT DOMAIN 2376 3176 NONHELICAL REGION.
FT SITE 2040 2042 CELL ATTACHMENT SITE.
FT SITE 2136 2138 CELL ATTACHMENT SITE.
FT SITE 2148 2150 CELL ATTACHMENT SITE.
FT SITE 2154 2156 CELL ATTACHMENT SITE.
FT SITE 2370 2372 CELL ATTACHMENT SITE.

Wed Jan 31 13:44:16 2001

```
FT DOMAIN 242 392 VWFA.
FT DOMAIN 837 986 VWFA.
FT DOMAIN 1233 1381 VWFA.
FT DOMAIN 1838 1994 VWFA.
FT DOMAIN 2402 2555 VWFA.
FT DOMAIN 2619 2786 VWFA.
FT DOMAIN 2987 3076 FIBRONECTIN TYPE-III.
FT DOMAIN 3107 3176 BPTI/KUNITZ INHIBITOR.
FT ACT_SITE 3121 3122 REACTIVE BOND.
FT DISULFID 3111 3161
FT DISULFID 3120 3144
FT DISULFID 3136 3157
FT CARBOHYD 202 202 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2079 2079 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2331 2331 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2558 2558 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2677 2677 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2861 2861 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 3036 3036 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPIC 32 236 MISSING (IN ISOFORM 2).
FT VARIANT 1679 1679 G -> E (IN BM).
FT VARIANT 2831 /FTid=VAR_001910.
FT CONFLICT 127 128 D -> H.
FT CONFLICT 137 137 /FTid=VAR_001911.
FT CONFLICT 2157 2157 QS -> AK (IN REF. 6).
SQ SEQUENCE 3176 AA; 343548 MW; 5DF82563229BB2DD CRC64; R -> L (IN REF. 6).
P -> R (IN REF. 4).
```

Query Match 32.5%; Score 44.5; DB 1; Length 3176;
Best Local Similarity 44.8%; Pred. No. 2.5e+02;
Matches 13; Conservative 2; Mismatches 5; Indels 9; Gaps 2;

QY 5 GPTPRAGRAALQKLRR-----RLTE 25
||| | ||| | |
Db 1103 GPTPNTG-AALEFVLRLNVSSAGSRITE 1130
||| | ||| | |

Search completed: January 30, 2001, 17:06:04
Job time: 1225 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:54:49 ; Search time 254.27 seconds
(without alignments)
11.985 Million cell updates/sec

Title: US-09-518-931-2_COPY_239_264

Perfect score: 137

Sequence: 1 PEGWGPTPRAGRAALQLKLRRLTEL 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL15:

- 1: sp-archaea:
- 2: sp-bacteria:
- 3: sp-fungi:
- 4: sp-human:
- 5: sp-invertebrate:
- 6: sp-mammal:
- 7: sp-mhc:
- 8: sp-organelle:
- 9: sp-phase:
- 10: sp-plant:
- 11: sp-rodent:
- 12: sp-virus:
- 13: sp-vertebrate:
- 14: sp-unclassified:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	137	100.0	300	4	095407
2	55	40.1	616	6	097507
3	53.5	39.1	1236	4	015398
4	52.5	38.3	656	4	076103
5	52.5	38.3	659	4	09NXXH9
6	51	37.2	112	2	09P9Q2
7	51	37.2	2481	12	09WJD3
8	50.5	36.9	395	2	09KZ51
9	50	36.5	311	2	09L122
10	50	36.5	680	2	09S2C9
11	50	36.5	789	11	09JLV2
12	49.5	36.1	200	2	069549
13	49	35.8	107	3	09P642
14	49	35.8	785	12	P89451
15	49	35.8	841	11	09WVL8
16	49	35.8	1592	10	09LGT4
17	48.5	35.4	798	5	09XTG1
18	48	35.0	183	5	09VRI3
19	48	35.0	258	4	09UF94

20	48	35.0	274	2	09RKB3
21	48	35.0	451	2	09RY44
22	48	35.0	602	4	09UJ72
23	48	35.0	1117	4	09ULB8
24	47.5	34.7	1703	11	09Z019
25	47	34.3	136	10	004355
26	47	34.3	160	2	09JZD8
27	47	34.3	409	5	09V756
28	47	34.3	495	2	09RZ27
29	47	34.3	504	4	075850
30	47	34.3	1044	4	09P206
31	46.5	33.9	232	7	P80943
32	46.5	33.9	333	6	028565
33	46.5	33.9	333	6	029422
34	46.5	33.9	658	2	09KNG3
35	46.5	33.9	750	10	09SDF5
36	46.5	33.9	1520	4	09Y4E0
37	46.5	33.9	2460	11	064512
38	46	33.6	155	5	023032
39	46	33.6	229	2	09X738
40	46	33.6	229	12	065978
41	46	33.6	449	4	014570
42	46	33.6	449	4	09NU86
43	46	33.6	506	2	050206
44	46	33.6	535	2	09XAV7
45	46	33.6	799	11	09Z268

ALIGNMENTS

RESULT 1

ID 095407 PRELIMINARY: PRT: 300 AA.

AC 095407;

DT 01-MAY-1999 (TREMBLrel. 10, Created)

DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)

DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)

DE DECOY RECEPTOR 3 (M68) (M68C) (M68E).

GN DCR3 OR TR6.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99087326; PubMed=9872321;

RA Pitti R.M., Marsters S.A., Lawrence D.A., Roy M., Kischkel F.C.,

RA Dowd P., Huang A., Donahue C.J., Sherwood S.W., Baldwin D.T.,

RA Godowski P.J., Wood W.I., Gurney A.L., Hillan K.J., Cohen R.L.,

RA Goddard A.D., Botstein D., Ashkenazi A.;

RT "Genomic amplification of a decoy receptor for Fas ligand in lung and

RT colon cancer.";

RL Nature 396:699-703(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=BLOOD;

RA MEDLINE=99253915; PubMed=10318773;

RY Yu K.Y., Kwon B., Ni J., Zhai Y., Ebner R., Kwon B.S.;

RT "A newly identified member of tumor necrosis factor receptor

superfamily (TR6) suppresses lIgT-mediated apoptosis.";

RL J. Biol. Chem. 274:13733-13736(1999).

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=PANCREAS;

RA MEDLINE=20126600; PubMed=10655513;

RY Bai C., Connolly B., Metzker M.L., Hilliard C.A., Liu X., Sandig V.,

RT "Overexpression of M68/DCR3 in human gastrointestinal tract tumors

independent of gene amplification and its location in a four-gene

cluster.";

RL Proc. Natl. Acad. Sci. U.S.A. 97:1230-1235(2000).

DR EMBL; AF104419; AAD03056.1; -.

Wed Jan 31 13:44:16 2001

```

DR EMBL; AF134240; AAD29688.1; -
DR EMBL; AF217796; AAF35244.1; -
DR EMBL; AF217793; AAF33685.1; -
DR EMBL; AF217794; AAF33686.1; -
DR HSSP; P25942; 1CDF.
DR INTERPRO; IPR000561; -
DR INTERPRO; IPR001368; -
DR PFAM; PF00020; TNFR_G6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS00050; TNFR_NGFR_2; 2.
DR PRODOM; PD00071; -; 1.
KW Receptor.
SQ SEQUENCE 300 AA; 32679 MW; F90AEE33718449AF CRC64;

Query Match 100.0%; Score 137; DB 4; Length 300;
Best Local Similarity 100.0%; Pred. No. 3.3e-12;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEGWGTTPRAGRAALQLKRLRRLTEL 26
DB 239 PEGWGTTPRAGRAALQLKRLRRLTEL 264

RESULT 2
O97507 PRELIMINARY; PRT; 616 AA.
AC O97507;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
EXII.
DE Sus scrofa (Pig).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
CX NCBI_TaxID=9823;
RN [1]
RP TISSUE=LIVER;
RA Takahashi T., Kihara T.;
RT "Porcine liver factor XII. ";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB022426; BAA37148.1; -
DR HSSP; P00750; 1KTF.
DR MEROPS; S01.211; -
DR INTERPRO; IPR000001; -
DR INTERPRO; IPR000083; -
DR INTERPRO; IPR000561; -
DR INTERPRO; IPR000562; -
DR INTERPRO; IPR000742; -
DR INTERPRO; IPR001254; -
DR INTERPRO; IPR001314; -
DR PFAM; PF00008; EGF; 2.
DR PFAM; PF00039; fn1; 1.
DR PFAM; PF00040; fn2; 1.
DR PFAM; PF00051; kringle; 1.
DR PFAM; PF00089; trypsin; 1.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR0018; CHYMOTRYPSIN.
DR PRINTS; PR00722; KRINGLE_1; UNKNOWN_1.
DR PROSITE; PS00021; KRINGLE_1; UNKNOWN_1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00134; TRYPsin_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PRODOM; PD00095; -; 1.
SQ SEQUENCE 616 AA; 68012 MW; 4C5FE3D71EBBD1A9 CRC64;

Query Match 40.1%; Score 55; DB 6; Length 616;
Best Local Similarity 46.2%; Pred. No. 6.2;
Matches 12; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 1 PEGWGTTPRAGRAALQLKRLRRLTEL 26
DB 344 PEGWGTTPRAGRAALQLKRLRRLTEL 369

RESULT 3
O15398 PRELIMINARY; PRT; 1236 AA.
ID O15398;
AC O15398;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE N-METHYL-D-ASPARTATE RECEPTOR 2C SUBUNIT PRECURSOR.
GN NMDAR2C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Daggett L.P., Lu C., Johnson E.C., Velicelebi G.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U77782; AAB71624.1; -
DR HSSP; P19491; 1GR2.
DR INTERPRO; IPR001311; -
DR INTERPRO; IPR001320; -
DR PFAM; PF00060; lig_chan; 1.
KW Signal.
FT SIGNAL 1 21
FT CHAIN 22 1236
FT SEQUENCE 1236 AA; 134462 MW; E57D38B26290A0F1 CRC64;
SQ SEQUENCE 1236 AA; 134462 MW; E57D38B26290A0F1 CRC64;

Query Match 39.1%; Score 53.5; DB 4; Length 1236;
Best Local Similarity 73.3%; Pred. No. 21;
Matches 11; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 PEGWGTTPRAGRAAL 15
DB 957 PTCWGP-PDGGRAAL 970

RESULT 4
O76103 PRELIMINARY; PRT; 556 AA.
ID O76103;
AC O76103;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE R29425.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E., McCready P.M., Skowronski E., Adamson A.W.,
RA Burkhardt-Schultz K., Gordon L., Kyle A., Ramirez M., Stilwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Barnes J.,
RA Dangnan L., Poundstone P., Christensen M., Georgescu A., Avila J.,
RA Liu S., Attix C., Andreise T., Frankheim M., Amico-Keller G.,
RA Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,
RA Krommiller B., Arellano A., Montgomery M., Ow D., Nolan M., Trong S.,
RA Kobayashi A., Olsen A.S., Carrano A.V.;
RT D193840.1;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005546; AAC33150.1; -
DR INTERPRO; IPR000051; -

```

DR INTERPRO: IPR000571; -
 DR INTERPRO: IPR002905; -
 DR PFAM: PF00642; zf-CCCH; 1.
 DR PFAM: PF02005; TRM; 1.
 SQ SEQUENCE 656 AA; 71787 MW; 510954A6073DF909 CRC64;

Query Match 38.3%; Score 52.5; DB 4; Length 656;
 Best Local Similarity 50.0%; Pred. No. 15;
 Matches 15; Conservative 5; Mismatches 5; Indels 5; Gaps 3;

QY 1 PE-GMGPTPRA---GRAALQ-LKLRRLTE 25
 II III III I:II : : : : :
 Db 552 PEANGPRPRARPGGKAADAEERRRLIQ 581

RESULT 5

Q9NXH9 PRELIMINARY; PRT: 659 AA.

AC Q9NXH9;

DT 01-OCT-2000 (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)

DE CDNA FLJ20244 FIS, CLONE COLF6450.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=COLON MUCOSA;

RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,

RA Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T.,

RA Nakamura Y., Isogai T., Sugano S.; -;

RT "NEDO human cDNA sequencing project."

RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK000251; BAA91031.1; -;

SQ SEQUENCE 659 AA; 72233 MW; E4F0F2B740B44387 CRC64;

Query Match

Best Local Similarity 38.3%; Score 52.5; DB 4; Length 659;

Matches 15; Conservative 5; Mismatches 5; Indels 5; Gaps 3;

QY 1 PE-GMGPTPRA---GRAALQ-LKLRRLTE 25
 II III III I:II : : : : :
 Db 555 PEANGPRPRARPGGKAADAEERRRLIQ 584

RESULT 6

Q9P9Q2 PRELIMINARY; PRT: 112 AA.

AC Q9P9Q2;

DT 01-OCT-2000 (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)

DE PHAGE-RELATED PROTEIN.

GN XF1678 OR XF1573.

OS Xylella fastidiosa.

OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;

OC Xylella.

OX NCBI_TaxID=2371;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=9ASC;

RX MEDLINE=20365717; PubMed=10910347;

RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,

RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,

RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,

RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carer H.,

RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,

RA Coutinho L.L., Cristofani M., Dias-Neto E., Dogena C., El-Dorry H.,

RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,

RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silva J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
 RA Vallada H., Van Sluys M.A., Verjovsky-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.; -;
 RT "The genome sequence of the plant pathogen Xylella fastidiosa";
 RL Nature 406:151-157(2000).
 DR EMBL; AE003993; AAF84487.1; -;
 DR EMBL; AE003996; AAF84382.1; -;
 SQ SEQUENCE 112 AA; 12762 MW; B8A8E83B4C8F4C70 CRC64;

Query Match 37.2%; Score 51; DB 2; Length 112;

Best Local Similarity 35.0%; Pred. No. 4.1;

Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 PEGWGPTPRAGRAALQLKLR 20
 I: I I I I I : : : : I
 Db 41 PDDWKPMPTVGRGVREIR 60

RESULT 7

Q9WJD3 PRELIMINARY; PRT: 2481 AA.

ID Q9WJD3;

DT 01-NOV-1999 (TREMBlrel. 12, Created)

DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)

DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)

DE NONSTRUCTURAL POLYPROTEIN.

OS Venezuelan equine encephalitis virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;

OC Chavivirus.

OX NCBI_TaxID=11036;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=AG80-663;

RX MEDLINE=99101297; PubMed=9886206;

RA Kinney R.M., Pfeiffer M., Tsuchiya K.R., Chang G.J., Roehrig J.T.;

RT Nucleotide sequences of the 26S mRNAs of the viruses defining the

RL Am. J. Trop. Med. Hyg. 59:952-964(1998).

DR EMBL; AF075258; AAD14564.1; -;

DR INTERPRO: IPR000606; -;

DR INTERPRO: IPR001788; -;

DR INTERPRO: IPR002589; -;

DR INTERPRO: IPR002620; -;

DR PFAM; PF00978; RNA_dep_RNAPol2; 2.

DR PFAM; PF01443; Viral_helicase1; 1.

DR PFAM; PF01661; DUF27; 1.

DR PFAM; PF01707; Peptidase_C9; 1.

KW Polyprotein.

FT CHAIN 1 535 METHYLTRANSFERASE NSP1.

FT CHAIN 536 1329 NONSTRUCTURAL PROTEINASE NSP2.

FT CHAIN 1330 1874 NONSTRUCTURAL PROTEIN NSP3.

FT CHAIN 1875 2481 RNA POLYMERASE NSP4.

SQ SEQUENCE 2481 AA; 275804 MW; 135F13F302E9695B CRC64;

Query Match

37.2%; Score 51; DB 12; Length 2481;

Wed Jan 31 13:44:16 2001

Best Local Similarity 52.6%; Pred. NO. 1e+02;
Matches 10; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 8 PRAGRAALQKLRLRLTEL 26
DB 453 PRSGSTALDIKLRDRIRKL 471

RESULT 8
Q9K251 PRELIMINARY; PRT; 395 AA.

AC O9K251;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE PUTATIVE PHOSPHOLIPASE.

GN SC3A4.07C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Brown S.P., Harris D.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Redenbach M., Kieser H.M., Denapaitte D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL354616; CAB89753.1; -;
SQ SEQUENCE 395 AA; 44541 MW; 7D2021762EEBB697 CRC64;

Query Match 36.9%; Score 50.5; DB 2; Length 395;
Best Local Similarity 57.1%; Pred. No. 18;
Matches 12; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

QY 6 PTPRAGRAALQKLRLRLTEL 26
DB 6 PTPGADRVSQIQLRLRLRL 25

RESULT 9
Q9L122 PRELIMINARY; PRT; 311 AA.

AC Q9L122;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE PUTATIVE LYSR-FAMILY TRANSCRIPTIONAL REGULATORY PROTEIN.
GN SC6D11.34.

OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);

RA Seeger K.J., Harris D.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);

RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Redenbach M., Kieser H.M., Denapaitte D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL158061; CAB76357.1; -;
SQ SEQUENCE 311 AA; 34370 MW; 581C4CAD64F8EA2D CRC64;

Query Match 36.5%; Score 50; DB 2; Length 311;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 12; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 3 GWGPTPRAGRAALQKLRLRLTEL 26
DB 55 GRLTPTAGARRLYPLQRLAEI 78

RESULT 10
Q9S2C9 PRELIMINARY; PRT; 680 AA.

AC Q9S2C9;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE PROBABLE SECRETED ALPHA-GALACTOSIDASE.

GN SCF85.12.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Seeger K.J., Harris D.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.

RC STRAIN=A3(2);
RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

RN [3]
RP SEQUENCE FROM N.A.

RC STRAIN=A3(2);
RA Redenbach M., Kieser H.M., Denapaitte D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).

DR EMBL; AL110470; CAB54169.1; -;

DR INTERPRO; IPR000111; -;

DR INTERPRO; IPR001917; -;

DR INTERPRO; IPR002241; -;

DR INTERPRO; IPR002287; -;

DR PFAM; PF02065; Melibiase; 2.

DR PRINTS; PR00740; GLHYDRILASE27.

DR PROSITE; PS00512; ALPHA-GALACTOSIDASE; 1.

DR PROSITE; PS00599; AA-TRANSFER_CLASS-2; UNKNOWN1.

SQ SEQUENCE 680 AA; 72167 MW; 7D63F1969E32BF5F CRC64;

Query Match 36.5%; Score 50; DB 2; Length 680;
Best Local Similarity 50.0%; Pred. No. 37;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 PEGWGPTRAGRAALQKL 18
DB 464 PDGWRVTPRSATSARALK 481

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:45:29 ; Search time 183.67 seconds
(without alignments)
3.910 Million cell updates/sec

Title: US-09-518-931-4_COPY_86_106

Perfect score: 124

Sequence: 1 ERCCRYCNVLCGEREEARACH 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_36.*

1: /SIDS1/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	124	100.0	153	20 Y22222	Human TNFR superfa
2	124	100.0	170	19 W63623	Human tumour necro
3	124	100.0	215	20 W93585	Human hAPO6 protei
4	124	100.0	245	20 Y28449	A human tumour nec
5	124	100.0	271	20 Y42184	Human mFLINT #1 pr
6	124	100.0	273	20 Y42185	Human mFLINT #2 pr
7	124	100.0	300	19 W66102	Amino acid sequenc
8	124	100.0	300	19 W63622	Human tumour.necro
9	124	100.0	300	20 Y03099	Human lung TNF-rec
10	124	100.0	300	20 Y42182	Human FLINT #1 pro
11	124	100.0	300	20 Y17479	Mammalian tumour n
12	124	100.0	300	20 Y06817	Human DCR3 polypep

13	124	100.0	300	20 W97749	Human tumour necro
14	124	100.0	300	20 W95082	Orphan receptor (H
15	124	100.0	302	20 Y42183	Human FLINT #2 pro
16	116	93.5	300	21 Y77458	Human TNF receptor
17	51	41.1	319	16 R74600	Follistatin. Synt
18	49	39.5	267	21 Y94890	Human protein clon
19	48	38.7	62	19 W4770	Fragment of scorpi
20	47	37.9	304	13 R20061	Pig EDF-binding pr
21	46.5	37.5	605	16 R74186	Chick p78. Gallus
22	46.5	37.5	605	21 Y76837	Netrin-1 protein s
23	46	37.1	127	20 Y41653	Zea mays XIAP asso
24	45	36.3	37	21 Y86345	Human gene 4-encod
25	45	36.3	122	17 W00499	Papillomavirus E2
26	45	36.3	202	21 Y86342	Human gene 4-encod
27	44	35.5	355	16 R85073	Cowpox virus T2-eq
28	44	35.5	3025	10 P93284	Sequence of clone
29	43.5	35.1	166	21 W90941	Pyrococcus sp. str
30	43.5	35.1	603	21 Y76839	Netrin protein seq
31	43.5	35.1	1539	21 Y4750	Human Smad Anchor
32	43	34.7	98	12 R14681	Fixx ferredoxin.
33	43	34.7	310	20 W87993	A human MCG4 prote
34	43	34.7	401	18 W38344	Mouse osteoprotege
35	42.5	34.3	604	18 W35946	Human netrin-1. H
36	42	33.9	78	20 Y30426	Mature nematode ex
37	42	33.9	106	17 R99947	Mutated OCIF, OCIF
38	42	33.9	143	17 R99946	Mutated OCIF, OCIF
39	42	33.9	145	17 R99930	Osteoclastogenesis
40	42	33.9	154	17 R99929	Osteoclastogenesis
41	42	33.9	162	20 Y30436	Mature nematode ex
42	42	33.9	181	17 R91711	AcanAP45. Ancylos
43	42	33.9	181	20 Y30409	Nematode extracted
44	42	33.9	187	17 R99950	Mutated OCIF, OCIF
45	42	33.9	187	21 Y77464	Primate protein se

ALIGNMENTS

RESULT 1
Y22222
ID Y22222 standard; Protein; 153 AA.
XX
AC Y22222;
XX
DT 16-SEP-1999 (first entry)
XX
DE Human TNFR superfamily soluble receptor protein sequence.
XX
KW TNFL1; human; TNFR superfamily; tumour necrosis factor ligand; TNF;
KW tumour necrosis factor receptor; TNFR superfamily; cell proliferation;
KW cell differentiation; cytokine production; immunoglobulin; hyperplasia;
KW apoptosis inducer; activated T cell; autoimmune disease; inhibitor;
KW myasthenia gravis; insulin-dependent diabetes mellitus; endotoxin shock;
KW rheumatoid arthritis; multiple sclerosis; systemic lupus erythematosus;
KW tumour; proliferative disorder; neoplasia; dysplasia; immunocompetence;
KW lymphoid organogenesis; bacterial resistance; contact hypersensitivity;
KW delayed type sensitivity; therapy.
XX
OS Homo sapiens.
XX
PN W09933980-A2.
XX
PD 08-JUL-1999.
XX
PF 22-DEC-1998; 98WO-US27474.
XX
PR 16-DEC-1998; 98US-0212270.
XX
PR 30-DEC-1997; 97US-0068959.
XX
(CHIR) CHIRON CORP.
XX
PI Kassam A, Lamson G, Pot D, Tribouley C;

3

XX New Tumor Necrosis Factor family receptor polypeptides and ligands -
 PT useful for diagnosis and treatment of prostate cancer and
 PT developmental or gestational abnormalities
 XX
 XX Claim 29; Fig 9; 156pp; English.
 XX
 CC This invention describes isolated Tumor Necrosis Factor (TNF) family
 CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
 CC their active fragments. APO4 is useful for diagnosing prostate cancer
 CC by determining levels of APO4 in an individual. Prostate cancer can also
 CC be treated using APO4 selective binding agents linked to a therapeutic
 CC moiety. APO4 polypeptides are also useful for identifying selective
 CC binding agents, useful in diagnosis/treatment of disease by binding of
 CC agents to the polypeptide/active fragment which is extracellular, or
 CC expressed on the cell surface. The binding is preferably performed in
 CC vivo. APO4 polypeptides/ active fragments are also useful for screening
 CC for agonists and antagonists by binding and observing the change in APO4
 CC activity. Effective pharmacological agents useful in diagnosis or
 CC treatment of disease are also identified using APO4 polypeptides/active
 CC fragments and APO4 signal transducer molecules that specifically interact
 CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4
 CC activity. The method is performed in vivo or in vitro. APO polypeptides
 CC are all useful as immunogens for preparing antibodies. APO4 is also
 CC useful for diagnosis/treatment of developmental or gestational
 CC abnormalities. APO8 was transfected to human breast carcinoma cell line
 CC MCF-7, and induced apoptosis.
 XX
 XX Sequence 215 AA;

Query Match 100.0%; Score 124; DB 20; Length 215;
 Best Local Similarity 100.0%; Pred. No. 4e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ERCRYCNVLCGEREEARACH 21
 Db 1 ercrycnvlgereearach 21
 |||||

RESULT 4
 Y28449
 ID Y28449 standard; Protein; 245 AA.
 XX
 AC Y28449;

XX 29-SEP-1999 (first entry)

XX A human tumour necrosis factor-R2-like proteins (TR2P)-1.

XX Human tumour necrosis factor-R2-like protein; TR2P; achondroplasia;
 KW osteoporosis; developmental disorder; Cushing's syndrome;
 KW muscular dystrophy; epilepsy; hereditary neuropathy;
 KW Charcot-Marie-Tooth disease; neurofibromatosis; hypothyroidism;
 KW hydrocephalus; seizure disorder; cerebral palsy; spinal bifida;
 KW congenital glaucoma; cataract; sensorineural hearing loss;
 KW reproductive disorder; infertility; ovulatory defect; endometriosis;
 KW autoimmune disorder; ectopic pregnancy; teratogenesis; spermatogenesis;
 KW immunological disorder; AIDS; Addison's disease; allergy; bronchitis;
 KW atherosclerosis; diabetes mellitus; Chron's disease; lupus;
 KW irritable bowel syndrome; multiple sclerosis; infection;
 KW neoplastic disorder; adenocarcinoma; leukaemia; lymphoma; melanoma;
 KW myeloma; sarcoma.

XX Homo sapiens.

OS WO9931128-A2.

XX 24-JUN-1999.

XX 02-DEC-1998; 98WO-US25649.

PR 16-DEC-1997; 97US-0991945.

XX (INCY-) INCYTE PHARM INC.

PI Au-Young J, Bandman O, Hillman JL, Kaser MR, Tang YT;

DR WPI; 1999-457916/38.

DR N-PSDB; X89503.

XX New tumour necrosis factor-R2-like protein - useful in the treatment
 PT of osteogenesis, developmental, reproductive, immunological and
 PT neoplastic disorders

PS Claim 1; Fig 1A-C; 81pp; English.

XX The present sequence represents a human tumour necrosis factor-R2-like
 CC protein (TR2P). The protein is used to treat and prevent osteogenesis,
 CC developmental, reproductive, immunological and neoplastic disorders, and
 CC also to diagnose disorders associated with TR2 protein expression. Such
 CC disorders include osteogenesis disorders such as achondroplasia and
 CC osteoporosis, developmental disorders such as Cushing's syndrome,
 CC muscular dystrophy, and epilepsy, hereditary neuropathies such as
 CC Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism,
 CC hydrocephalus, seizure disorders such as cerebral palsy and spinal
 CC bifida, congenital glaucoma, cataract, or sensorineural hearing loss,
 CC reproductive disorders such as infertility, ovulatory defects and
 CC endometriosis, autoimmune disorders, ectopic pregnancy and teratogenesis,
 CC disruption of spermatogenesis, immunological disorders such as AIDS,
 CC Addison's disease, allergies, bronchitis, atherosclerosis, diabetes
 CC mellitus, Chron's disease, lupus and irritable bowel syndrome, multiple
 CC sclerosis, viral, fungal, helminthic, parasitic and protozoal infections,
 CC and neoplastic disorders including adenocarcinoma, leukaemia, lymphoma,
 CC melanoma, myeloma, sarcoma, and teratocarcinoma.

XX Sequence 245 AA;

Query Match 100.0%; Score 124; DB 20; Length 245;
 Best Local Similarity 100.0%; Pred. No. 4.5e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ERCRYCNVLCGEREEARACH 21
 Db 86 ercrycnvlgereearach 106
 |||||

RESULT 5
 Y42184
 ID Y42184 standard; Protein; 271 AA.
 XX
 AC Y42184;

XX 17-DEC-1999 (first entry)

XX Human mFLINT #1 protein sequence.

XX Human; FLINT; mFLINT; OP63; tumour necrosis factor receptor; FasL;
 KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
 KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
 KW reperfusion-associated injury; aplastic anaemia; differentiation;
 KW growth; myelodysplastic syndrome; pancytopenic condition;
 KW myocardial ischaemia.

XX Homo sapiens.

XX WO9950413-A2.

XX 07-OCT-1999.

XX 30-MAR-1999; 99WO-US06797.

XX 30-MAR-1998; 98US-0079856.

PR 20-MAY-1998; 98US-0086074.

Wed Jan 31 13:45:15 2001

```

PR 09-SEP-1998; 98US-0099643.
PR 17-DEC-1998; 98US-0112577.
PR 18-DEC-1998; 98US-0112703.
PR 18-DEC-1998; 98US-0112933.
PR 22-DEC-1998; 98US-0113407.
XX (ELIL ) LILLY & CO ELI.
PA Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
PI Song HY, Wang J, Wu X, Zuckerman SH;
XX WPI: 1999-591319/50.
DR N-PSDB: 225377.
XX
XX Use of mature FLINT for treating acute liver failure, inflammation,
PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
PT and proinflammatory activity
XX
XX Claim 31; Fig 3; 99pp; English.
PS
XX The present invention describes therapeutic applications of mature FLINT
CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
CC (mFLINT), which is a member of the tumour necrosis factor receptor
CC superfamily, is used for treating acute liver failure, inflammation of
CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
CC injury or disorder such as hypercoagulation (including use with
CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
CC or disorder, Type I diabetes, cancer, cell damage or damage to an
CC innocent bystander tissue that is induced by a chemotherapeutic agent or
CC therapeutic irradiation, treating haematopoietic progenitor cells that
CC have been exposed to therapeutic radiation or chemotherapy, aplastic
CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
CC also used for promoting the growth or differentiation of a haematopoietic
CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
CC resulting from abnormal myocardial ischaemia. The present sequence
XX represents human mFLINT.
XX
XX Sequence 271 AA;
SQ
Query Match 100.0%; Score 124; DB 20; Length 271;
Best Local Similarity 100.0%; Pred. No. 5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ECRYCNVLCGEREEARACH 21
DB 57 ERCRCNVLCGEREEARACH 77

RESULT 6
Y42185
ID Y42185 standard; Protein; 273 AA.
XX
XX Y42185;
AC
XX
XX 17-DEC-1999 (first entry)
XX
XX Human mFLINT #2 protein sequence.
XX
XX Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
KW reperfusion-associated injury; aplastic anaemia; differentiation;
KW growth; myelodysplastic syndrome; pancytopenic condition;
KW myocardial ischaemia.
XX
XX Homo sapiens.
XX
XX WO950413-A2.
XX
XX 07-OCT-1999.
PD

```

```

XX 30-MAR-1999; 99WO-US06797.
XX
XX 30-MAR-1998; 98US-0079856.
PR 20-MAY-1998; 98US-0086074.
PR 09-SEP-1998; 98US-0099643.
PR 17-DEC-1998; 98US-0112577.
PR 18-DEC-1998; 98US-0112703.
PR 22-DEC-1998; 98US-0112933.
XX (ELIL ) LILLY & CO ELI.
XX
XX Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
PI Song HY, Wang J, Wu X, Zuckerman SH;
XX WPI: 1999-591319/50.
DR N-PSDB: 225378.
XX
XX Use of mature FLINT for treating acute liver failure, inflammation,
PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
PT and proinflammatory activity
XX
XX Example 2; Fig 4; 99pp; English.
PS
XX The present invention describes therapeutic applications of mature FLINT
CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
CC (mFLINT), which is a member of the tumour necrosis factor receptor
CC superfamily, is used for treating acute liver failure, inflammation of
CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
CC injury or disorder such as hypercoagulation (including use with
CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
CC or disorder, Type I diabetes, cancer, cell damage or damage to an
CC innocent bystander tissue that is induced by a chemotherapeutic agent or
CC therapeutic irradiation, treating haematopoietic progenitor cells that
CC have been exposed to therapeutic radiation or chemotherapy, aplastic
CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
CC also used for promoting the growth or differentiation of a haematopoietic
CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
CC resulting from abnormal myocardial ischaemia. The present sequence
XX represents human mFLINT.
XX
XX Sequence 273 AA;
SQ
Query Match 100.0%; Score 124; DB 20; Length 273;
Best Local Similarity 100.0%; Pred. No. 5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ECRYCNVLCGEREEARACH 21
DB 57 ERCRCNVLCGEREEARACH 77

RESULT 7
W66102
ID W66102 standard; Protein; 300 AA.
XX
XX W66102;
AC
XX
XX 02-DEC-1998 (first entry)
XX
XX Amino acid sequence of tumour necrosis related receptor (TR4).
XX
XX Human; tumour necrosis related receptor; TR4; agonist; antagonist;
KW inhibition; chronic; acute; inflammation; arthritis; septicaemia;
KW autoimmune disease; transplant rejection; stroke; cancer;
KW Alzheimer's disease.
XX
XX Homo sapiens.
XX
XX OS
XX
XX

```

```

PN EP861850-A1.
XX
PD 02-SEP-1998.
XX
PF 20-JAN-1998; 98EP-0300382.
XX
PR 04-FEB-1997; 97US-0794796.
XX
PA (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI Emery J, Tan KB, Truneh A, Young PR;
XX
DR WPI; 1998-508248/44.
XX
DR N-PSDB; V07654.
XX
XX New DNA encoding tumour necrosis related receptor - used to treat
PT and prevent e.g. inflammation, arthritis, septicaemia, autoimmune
PT diseases, transplant rejection, infection, stroke, ischaemia, ARDS,
PT restenosis, AIDS, bone disorders and cancer
XX
PS Claim 1; Fig 1; 21pp; English.
XX
XX This is the amino acid sequence of the human tumour necrosis related
CC receptor (TR4), used in the method of the invention. The TR4 protein
CC or its agonist can be used to treat a subject in need of enhanced
CC TR4 polypeptide activity. The antagonist is used to inhibit TR4
CC polypeptide activity. The active agents can be used for the
CC treatment and prevention of diseases such as chronic and acute
CC inflammation, arthritis, septicaemia, autoimmune diseases, transplant
CC rejection, stroke, cancer, Alzheimer's disease.
XX
SQ Sequence 300 AA;

Query Match 100.0%; Score 124; DB 19; Length 300;
Best Local Similarity 100.0%; Pred. No. 5.5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERCRYCNVLCGEREEARACH 21
Db 86 ercrycnvlgereearach 106

RESULT 8
W63622
ID W63622 standard; Protein; 300 AA.
XX
AC W63622;
XX
DT 26-OCT-1998 (first entry)
XX
DE Human tumour necrosis factor receptor-6 alpha protein.
XX
KW Human tumour necrosis factor receptor-6 alpha; TNFR-6 alpha; TNFR-6 beta;
KW endothelial cells; keratinocytes; normal prostate; apoptosis;
KW prostate tumour tissue.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..30
FT Protein 31..300
FT /note= "TNFR-6 alpha"
FT Region 31..282
FT /note= "Soluble extracellular domain"
XX
XX W09830694-A2.
PN
XX
PD 16-JUL-1998.
XX
PF 13-JAN-1998; 98WO-US00153.
XX
PR 14-JAN-1997; 97US-0035496.
XX

(PHUMA-) HUMAN GENOME SCI INC.
XX
PI Ebner R, Feng P, Gentz RL, Ni J, Ruben SM, Yu G;
XX
DR WPI; 1998-399142/34.
XX
DR N-PSDB; V39085.
XX
PT Human tumour necrosis factor receptors 6-alpha and 6-beta - used in
PT the diagnosis of immune system-related disorder(s)
XX
XX Claim 20; Fig 1; 91pp; English.
XX
XX The present sequence represents the human tumour necrosis factor
CC receptor-6 alpha (TNFR-6 alpha) protein. The invention also provides
CC for the TNFR-6 beta protein (W63623). TNFR-6 alpha and TNFR-6 beta are
CC members of the tumour necrosis factor receptor (TNFR) family. TNFRs
CC are expressed in endothelial cells, keratinocytes, normal prostate and
CC prostate tumour tissue. For a number of disorders of these cells,
CC particularly of the immune system, substantially altered (whether
CC increased or decreased) levels of TNFR-6 alpha and/or TNFR-6 beta gene
CC expression can be detected, therefore the TNFR-6 alpha and TNFR-6 beta
CC polypeptides, nucleic acids and antibodies are claimed to be useful in
CC the diagnosis of such disorders. Mutations of the TNFR-6 alpha and
CC TNFR-6 beta genes can also be detected. The TNFR polypeptides are
CC also claimed to be useful for identifying ligands which may be useful
CC in the treatment of apoptosis related disorders.
XX
SQ Sequence 300 AA;

Query Match 100.0%; Score 124; DB 19; Length 300;
Best Local Similarity 100.0%; Pred. No. 5.5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERCRYCNVLCGEREEARACH 21
Db 86 ercrycnvlgereearach 106

RESULT 9
Y03099
ID Y03099 standard; Protein; 300 AA.
XX
AC Y03099;
XX
DT 09-DEC-1999 (first entry)
XX
DE Human lung TNF-receptor protein.
XX
KW Tumour necrosis factor; TNF; TNF receptor; human; lung; gene therapy;
KW detection; immunoassay; diagnosis; disease; immune system; tumour;
KW osteogenic system; cardiovascular system; central nervous system; asthma;
KW peripheral nervous systems; transplant incompatibility; antitumor;
KW rheumatoid arthritis; antiasthmatic; antiarthritic.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 134..1036
FT /*tag= a
FT /product= "TNF-receptor"
XX
XX DE19809978-A1.
PN
XX
PD 16-SEP-1999.
XX
PF 09-MAR-1998; 98DE-1009978.
XX
PR 09-MAR-1998; 98DE-1009978.
XX
PA (BADI ) BASF AG.
XX

```

Wed Jan 31 13:45:15 2001

```

PI Kroeger B;
XX WPI; 1999-519473/44.
DR N-PSDB; Z09998.
XX
XX New soluble member of tumor necrosis factor receptor family, useful for
XX identification specific modulators and for treating disease e.g. tumors
XX
XX
XX Claim 1; Page 8-9; 10pp; German.
XX
XX This invention describes a novel tumour necrosis factor (TNF) receptor
XX (I) isolated from human lung tissue. (I) is used: (i) to raise specific
XX antibodies (Ab); (ii) to screen for specific (ant)agonists or ligands
XX (A), potential therapeutic agents; and (iii) therapeutically (optionally
XX expressed from a gene therapy vector) in conditions associated with a
XX deficit of (I). Ab are used: (a) for qualitative or quantitative
XX detection of (I) in standard immunoassays (for diagnosis of disease, or
XX susceptibility, or for monitoring); and (b) as therapeutic inhibitors in
XX cases where (I) is overexpressed. Nucleic acid (II) that encodes (I) is
XX used: (A) for recombinant production of (I); (B) also its oligonucleotide
XX fragments, in standard hybridization and/or amplification assays; (C) as
XX source of antisense molecules or ribozymes; and (D) to produce transgenic
XX animals (for studying (patho)physiology of (I)). Diseases possibly
XX associated with under- or over-expression of (I) are those of the immune,
XX osteogenic, cardiovascular and central or peripheral nervous systems, rhe
XX tumors, transplant incompatibility, asthma and rheumatoid arthritis. The
XX products of the invention have antitumor, antiasthmatic and
XX antiarthritic activity. This sequence represents the TNF-receptor of the
XX invention.
XX
XX Sequence 300 AA;
XX
XX Query Match 100.0%; Score 124; DB 20; Length 300;
XX Best Local Similarity 100.0%; Pred. No. 5.5e-09;
XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 ECRYCNVLCGEREEERACH 21
XX Db 86 ercrycnvlgereearach 106
XX
XX RESULT 10
XX Y42182
XX ID Y42182 standard; Protein; 300 AA.
XX AC Y42182;
XX DT 17-DEC-1999 (first entry)
XX DE Human FLINT #1 protein sequence.
XX
XX Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
XX apoptosis; inflammation; cancer; diabetes; acute liver failure;
XX sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
XX reperfusion-associated injury; aplastic anaemia; differentiation;
XX growth; myelodysplastic syndrome; pancytopenic condition;
XX myocardial ischaemia.
XX
XX Homo sapiens.
XX
XX OS WO9950413-A2.
XX PN 07-OCT-1999.
XX
XX 30-MAR-1999; 99WO-US06797.
XX
XX 30-MAR-1998; 98US-0079856.
XX PR 20-MAY-1998; 98US-0086074.
XX PR 09-SEP-1998; 98US-0099643.
XX PR 17-DEC-1998; 98US-0112577.
XX PR 18-DEC-1998; 98US-0112703.
XX

```

```

PR 18-DEC-1998; 98US-0112933.
PR 22-DEC-1998; 98US-0113407.
XX
XX (ELIL ) LILLY & CO ELI.
XX
XX Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JC;
XX Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
XX Song HY, Wang J, Wu X, Zuckerman SH;
XX WPI; 1999-591319/50.
XX DR N-PSDB; Z25375.
XX
XX Use of mature FLINT for treating acute liver failure, inflammation,
XX cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
XX and proinflammatory activity
XX
XX Claim 30; Fig 1; 99pp; English.
XX
XX The present invention describes therapeutic applications of mature FLINT
XX (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
XX (mFLINT), which is a member of the tumour necrosis factor receptor
XX superfamily, is used for treating acute liver failure, inflammation of
XX the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
XX with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
XX injury or disorder such as hypercoagulation (including use with
XX thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
XX or disorder, Type I diabetes, cancer, cell damage or damage to an
XX innocent bystander tissue that is induced by a chemotherapeutic agent or
XX therapeutic irradiation, treating haematopoietic progenitor cells that
XX have been exposed to therapeutic radiation or chemotherapy, aplastic
XX anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
XX also used for promoting the growth or differentiation of a cardiac myocyte
XX progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
XX resulting from abnormal myocardial ischaemia. The present sequence
XX represents human FLINT.
XX
XX Sequence 300 AA;
XX
XX Query Match 100.0%; Score 124; DB 20; Length 300;
XX Best Local Similarity 100.0%; Pred. No. 5.5e-09;
XX Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 ECRYCNVLCGEREEERACH 21
XX Db 86 ercrycnvlgereearach 106
XX
XX RESULT 11
XX Y17479
XX ID Y17479 standard; Protein; 300 AA.
XX AC Y17479;
XX
XX 02-AUG-1999 (first entry)
XX
XX Mammalian tumour necrosis factor receptor OPG-2.
XX
XX Tumour necrosis factor receptor; TNF receptor; OPG-2; Paget's disease;
XX osteopenic disorder; osteoclast activity; primary osteoporosis;
XX hyperglycaemia; osteolytic metastasis; immune response; cancer.
XX
XX Mammalia.
XX
XX OS WO9926977-A1.
XX PN 03-JUN-1999.
XX
XX 24-NOV-1998; 98WO-US25065.
XX
XX 17-FEB-1998; 98US-0074896.
XX PR 24-NOV-1997; 97US-0066446.
XX

```

RA Pauley A., Gattung S.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA waterston R.;
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U64854; AAB18318.1; -;
 DR INTERPRO: IPR000107; -;
 DR INTERPRO: IPR000059; -;
 DR INTERPRO: IPR001682; -;
 DR INTERPRO: IPR002048; -;
 DR INTERPRO: IPR002202; -;
 DR INTERPRO: IPR003032; -;
 DR PFAM; PF000036; efhand; 1.
 DR PFAM; PF00622; SPRY; 3
 DR PFAM; PF01365; RYDR_ITPR; 2.
 DR PFAM; PF02026; RYR; 4.
 DR PROSITE; PS00018; EF_HAND; UNKNOWN_2.
 DR PROSITE; PS00066; HMG_COA_REDUCTASE_1; UNKNOWN_1.
 SQ SEQUENCE 5107 AA; 580731 MW; 462b7c01b799d152 CRC64;

Query Match 49.5%; Score 45.5; DB 5; Length 5107;
 Best Local Similarity 71.4%; Pred. No. 2.8e+02;
 Matches 10; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 2 ETPIYPRDAETGE 15
 I | | | | : | | | | |
 DB 2414 ENPTYPSKDAE-GE 2426

RESULT 5
 ID Q9QY29 PRELIMINARY; PRT; 310 AA.
 AC Q9QY29;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE DISTAL INTESTINAL SERINE PROTEASE.
 GN DISP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Shaw-Smith C.J., Coffey A.J., Leversha M., Freeman T.C., Bentley D.R.,
 RA Walters J.R.;
 RT "Characterization of a novel murine intestinal serine protease,
 RT DISP."
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ243866; CAB56465.1; -;
 DR HSSP; P20231; 1AAO.
 DR INTERPRO: IPR001254; -;
 DR INTERPRO: IPR001314; -;
 DR PFAM; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Protease.
 SQ SEQUENCE 310 AA; 33701 MW; F828EC7F6D25303F CRC64;

Query Match 48.9%; Score 45; DB 11; Length 310;
 Best Local Similarity 58.3%; Pred. No. 17;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 PTYPWRDAETGE 15
 I | | | | : | | | | |
 DB 117 PTYLWADASSGD 128

RESULT 6
 ID Q9L8C7 PRELIMINARY; PRT; 7257 AA.
 AC Q9L8C7;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE POLYKETIDE SYNTHASE.
 GN EPOC.
 OS Sorangium cellulosum.
 OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
 OC Myxococcales; Sorangineae; Sorangiales; Sorangium.
 OX NCBI_TaxID=39644;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SO CE90;
 RA Molnar I., Schupp T., Ono M., Zirkle R.E., Milnamow M.,
 RA Nowak-Thompson B., Engel N., Toupet C., Stratmann A., Cyr D.D.,
 RA Grolach J., Mayo J.M., Hu A., Goff S., Schmid J., Ligon J.M.;
 RT "The biosynthetic gene cluster for the microtubule-stabilizing agents
 RT epothilones A and B from Sorangium cellulosum So ce90.";
 RL Chem. Biol. 7:97-109(2000).
 DR EMBL; AF210843; AAF26921.1; -;
 SQ SEQUENCE 7257 AA; 764299 MW; 621A6820FA57ACE9 CRC64;

Query Match 48.9%; Score 45; DB 2; Length 7257;
 Best Local Similarity 52.9%; Pred. No. 4.8e+02;
 Matches 9; Conservative 1; Mismatches 1; Indels 6; Gaps 1;

QY 4 PTYPWR-----DAETG 14
 I | | | | : | | | | |
 DB 5968 PTYPQWRERYVVDATG 5984

RESULT 7
 ID Q9W0P8 PRELIMINARY; PRT; 1521 AA.
 AC Q9W0P8;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE CG12031 PROTEIN.
 GN CG12031.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yeandle M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Balley R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foshier C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

Q9NKB5	PRELIMINARY;	PRT: 1584 AA.		
AC				
DT	01-OCT-2000 (TrEMBLrel. 15, Created)			
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)			
DT	01-OCT-2000 (TrEMBLrel. 15, Last annotation update)			
DE	HYPOPHETICAL 183.0 KDA PROTEIN.			
GN	BG:DS07721.6.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RC	SEQUENCE FROM N.A.			
RP	STRAIN=Y, CN BW SP;			
RX	MEDLINE=99403001; PubMed=10471707;			
RA	Ashburner M., Mista S., Roote J., Lewis S.E., Blazej R., Davis T.,			
RA	Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,			
RA	Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,			
RA	Palazzo M., Reese M.G., Spradling A., Tsang K., Wan K., Whitehead K.,			
RA	Celniker S., Rubin G.M.;			
RT	"An exploration of the sequence of a 2.9-Mb region of the genome of			
RT	Drosophila melanogaster: the Adh region.";			
RL	Genetics 153:179-219(1999).			
RN	[2]			
RC	SEQUENCE FROM N.A.			
RP	STRAIN=Y, CN BW SP;			
RA	Celniker S.E., Agbayani A., Arcaina T.T., Baxter E., Blazej R.G.,			
RA	Butenhoff C., Champagne M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,			
RA	Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,			
RA	Houston K.A., Hummasti S.R., Karra K., Kearney L., Kim E., Lee B.,			
RA	Lewis S., Li P., Lomutan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,			
RA	Nixon K., Pachter J.M., Park S., Pfeiffer B., Poon L., Sequelra A.,			
RA	Sethi H., Snier E., Svirskas R.R., Wan K.H., Weinburg T., Zhang R.,			
RA	Zieran L.L., Rubin G.M.;			
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AE003410; AAF44887.1; -;			
DR	FLYBASE; Fbgn0028854; BG:DS07721.			
KW	Hypothetical protein.			
SQ	SEQUENCE 1584 AA; 183014 MW; 5953A0FC8B77F47D CRC64;			
Query Match 46.2%; Score 42.5; DB 5; Length 1584;				
Best Local Similarity 42.1%; Pred No. 2.4e+02;				
Matches 8; Conservative 4; Mismatches 2; Indels 5; Gaps				
Qy	2 ETP-----TYPWRDAETGE 15			
Db	: : : :			
153 ESPEGGTSHPWRSAGK 171				
RESULT 14				
Q9NNP0	PRELIMINARY; PRT: 338 AA.			
AC	Q9NNP0			
DT	01-OCT-2000 (TrEMBLrel. 15, Created)			
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)			
DT	01-OCT-2000 (TrEMBLrel. 15, Last annotation update)			
DE	POSSIBLE PODOCALYXIN-LIKE PROTEIN 1 (FRAGMENT).			
GN	LM15.94.			
OS	Leishmania major.			
OC	Eukaryota; Eudlenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.			
OX	NCBI_TaxID=5664;			
RN	[1]			
RC	SEQUENCE FROM N.A.			
RP	STRAIN=FRIDLIN;			
RA	Murphy L., Quail M., Harris D., Rajandream M., Ivens A., Barrell B.;			
RL	Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AL160371; CAC00170.1; -;			
FT	NON_TER 1			
FT	NON_TER 338			
SQ	SEQUENCE 338 AA; 33865 MW; A5B227D9ACF6059 CRC64;			

Wed Jan 31 13:44:31 2001

Query Match 45.7%; Score 42; DB 5; Length 338;
Best Local Similarity 56.2%; Pred. No. 56;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 AETPTPWDAETGER 16
| | | | : | | | |
Db 149 AATPTGPGDGENGER 164

RESULT 15

Q9THU1 PRELIMINARY; PRT; 475 AA.
AC Q9THU1;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE RIBULOSE-1,5-BISPHOSPHATE CARBOXYLASE/OXYGENASE, LARGE SUBUNIT
(FRAGMENT).
DE RBCL.
GN RBCL.
OS Kalmia polifolia.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; Ericales;
OC Ericaceae; Kalmia.
OX NCBI_TaxID=45899;
RN [1]
RP SEQUENCE FROM N.A.
RA Kron K.A.; King J.M.;
RT "Cladistic relationships of Kalmia, Leiothylum, and Loiseleuria."
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U49289; AAF14329.1; -
DR HSSP; P00880; 1RBL.
DR INTERPRO; IPR000685; -
DR PFAM; PF00016; RUBISCO_large; 1.
DR PROSITE; PS00157; RUBISCO_LARGE; 1.
KW Chloroplast.
FT NON_TER 1
SQ SEQUENCE 475 AA; 52415 MW; 91D42E8A372F0816 CRC64;

Query Match 45.1%; Score 41.5; DB 8; Length 475;
Best Local Similarity 40.0%; Pred. No. 97;
Matches 10; Conservative 0; Mismatches 2; Indels 13; Gaps 1;

QY 4 PTTPWRD-----AETGE 15
| | | | |
Db 210 PFIPWRDRFLFCAEALYKQAETGE 234

Search completed: January 30, 2001, 16:54:41
Job time: 897 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:47:40 ; Search time 132.77 Seconds
(without alignments)
2.164 Million cell updates/sec

Title: US-09-518-931-2_COPY_31_46
Perfect score: 92
Sequence: 1 AETPTYPWRDAETGER 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	92	100.0	300	2	US-08-794-796-2
2	47	51.1	775	2	US-08-966-388-4
3	47	51.1	775	3	US-09-188-403-4
4	47	51.1	775	3	US-09-188-404-4
5	45	48.9	7257	3	US-09-335-409-5
6	43.5	47.3	1421	3	US-09-335-409-2
7	41	44.6	797	2	US-08-663-566A-2
8	41	44.6	797	2	US-08-023-610-2
9	41	44.6	797	2	US-08-288-065A-2
10	41	44.6	797	2	US-08-362-240A-2
11	41	44.6	797	4	PCT-US95-10245-2
12	40	43.5	268	1	US-08-440-103-29
13	40	43.5	268	1	US-08-440-103-30
14	40	43.5	268	1	US-08-440-542-29
15	40	43.5	268	1	US-08-440-542-30
16	40	43.5	268	1	US-08-231-368-29
17	40	43.5	268	1	US-08-231-368-30
18	40	43.5	268	1	US-08-440-210-29
19	40	43.5	268	1	US-08-440-210-30
20	40	43.5	269	1	US-08-440-103-16
21	40	43.5	269	1	US-08-440-542-16
22	40	43.5	269	1	US-08-231-368-16
23	40	43.5	269	1	US-08-440-210-16
24	40	43.5	402	1	US-08-460-806-2
25	40	43.5	402	1	US-08-325-630-2
26	39	42.4	932	3	US-08-968-752B-6
27	39	42.4	1003	3	US-08-851-843A-217
28	39	42.4	1003	3	US-08-974-549A-336

29	39	42.4	1156	3	US-08-996-083-1	Sequence 1, Appli
30	38	41.3	242	1	US-08-289-699A-6	Sequence 6, Appli
31	38	41.3	242	2	US-08-878-283-6	Sequence 6, Appli
32	38	41.3	244	1	US-08-289-699A-3	Sequence 3, Appli
33	38	41.3	244	2	US-08-878-283-3	Sequence 3, Appli
34	38	41.3	1832	3	US-09-335-409-4	Sequence 4, Appli
35	38	41.3	2439	3	US-09-335-409-7	Sequence 7, Appli
36	38	41.3	3798	3	US-09-335-409-6	Sequence 6, Appli
37	37	40.2	402	1	US-08-460-806-9	Sequence 9, Appli
38	37	40.2	402	1	US-08-460-806-11	Sequence 11, Appli
39	37	40.2	402	1	US-08-325-630-9	Sequence 9, Appli
40	37	40.2	402	1	US-08-325-630-11	Sequence 11, Appli
41	37	40.2	480	1	US-08-440-103-18	Sequence 18, Appli
42	37	40.2	480	1	US-08-440-542-18	Sequence 18, Appli
43	37	40.2	480	1	US-08-231-368-18	Sequence 18, Appli
44	37	40.2	480	1	US-08-440-210-18	Sequence 18, Appli
45	37	40.2	508	1	US-07-891-942G-10	Sequence 10, Appli

ALIGNMENTS

RESULT 1
US-08-794-796-2
; Sequence 2, Application US/08794796
; Patent No. 5885800
; GENERAL INFORMATION:
; APPLICANT: Emery, John
; APPLICANT: Tan, KB
; APPLICANT: Truneh, Alem
; APPLICANT: Young, Peter
; TITLE OF INVENTION: Tumor Necrosis Related Receptor,
; TITLE OF INVENTION: TR4
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithkline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/794,796
; FILING DATE: 04-FEB-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: GH50000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-4026
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 300 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-794-796-2

Query Match 100.0%; Score 92; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;

16. Conservative	0; Mismatches	0; Indels	0; Gaps
------------------	---------------	-----------	---------

QY 1 AETPTYPWRDAETGER 16
| | | | | | | | | | | | | | | |
DB 31 AETPTYPWRDAETGER 46

RESULT 2
US-08-966-388-4
; Sequence 4, Application US/08966388
; Patent No. 5965412
; GENERAL INFORMATION:
; APPLICANT: TOMOYUKI NISHIMOTO
; APPLICANT: HIROTO KUBOTA
; APPLICANT: HIROTO CHAEN
; APPLICANT: TOSHIO MIYAKE
; ITS PREPARATION AND USES

;
; CORRESPONDENCE ADDRESS:
; ADDRESS: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
;

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

CLASSIFICATION: 435
 PRIOR APPLICATION DATA: JP 311,235/1996
 FILING DATE: 8-NOV-1996
 PRIOR APPLICATION DATA: JP 61,710/97
 FILING DATE: 3-MAR-1997
 ATTORNEY/AGENT INFORMATION:

; NAME: BROWDY, Roger L.
 ; REGISTRATION NUMBER: 25,618
 ; REFERENCE/DOCKET NUMBER:
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-628-5197
 ; TELEFAX: 202-737-3528
 ; TELEX: 248633
 ;

```

; TELE: 248033
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 775 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-966-388-4

```

Query Match 51.1%; Score 47; DB 2; Length 775;
Best Local Similarity 70.0%; Pred. No. 8.7;
Matches 7; Conservative 1; Mismatches 2; Indels

Qy 6 YPWRDAETGE 15
||| | : |||
Db 407 YPWESADTGE 416

RESULT 3
US-09-188-403-4
; Sequence 4, Application US/09188403
; Patent No. 6066477
; GENERAL INFORMATION:
; APPLICANT: TOMOYUKI NISHIMOTO
; APPLICANT: MICHIO KUBOTA

APPLICANT: HIROTO CHAEN
APPLICANT: TOSHIO MIYAKE
TITLE OF INVENTION: KJOTIBOSE PHOSPHORYLASE, ITS PREPARATION AND USES
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/188.403
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/966.388
FILING DATE:

PRIOR APPLICATION DATA: JP 61,710/97
APPLICATION NUMBER: 3-NAR-1997
FILING DATE: 3-NAR-1997
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/POCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEFAX: 2486333

```

      / SELEX: ZR003
      / INFORMATION FOR SEQ ID NO: 4:
      / SEQUENCE CHARACTERISTICS:
      /   LENGTH: 775 amino acids
      /   TYPE: amino acid
      /   TOPOLOGY: linear
      / MOLECULE TYPE: peptide
      / US-09-188-403-4

```

Query Match 51.1%; Score 47; DB 3; Length 775;
Best Local Similarity 70.0%; Pred. No. 8.7;
Matches 7: Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 6 YPWRDAETGE 15
||| | : |||
Dh 407 YPWESADTGE 416

RESULT 4
US-09-188-404-4
; Sequence 4, Application US/09188404
; Patent No. 6140487
; GENERAL INFORMATION:
; APPLICANT: TOMOVUKI NISHIMOTO

APPLICANT: TOMOYUKI NISHIMOTO
APPLICANT: Michio KUBOTA
APPLICANT: Hiroto CHAEN
APPLICANT: Toshio MIYAKE
TITLE OF INVENTION: KOJIBIOSE PHOSPHORYLASE, ITS PREPARATION AND USES
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/188,404
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: 08/966,388
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 61,710/97
; FILING DATE: 3-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 775 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-188-404-4

Query Match 51.1%; Score 47; DB 3; Length 775;
Best Local Similarity 70.0%; Pred. No. 8.7;
Matches 7; Conservative 1; Mismatches 2; Indels 2; Gaps 0;

Qy 6 YPWRDAETGE 15
||| |:
Db 407 YPWESADTG 416

RESULT 5
US-09-335-409-5
; Sequence 5, Application US/09335409
; Patent No. 6121029
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/335,409
; CURRENT FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 7257
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
; US-09-335-409-5

Query Match 48.9%; Score 45; DB 3; Length 7257;
Best Local Similarity 52.9%; Pred. No. 2.2e+02;
Matches 9; Conservative 1; Mismatches 1; Indels 6; Gaps 1;

Qy 4 PTYPWR-----DAETG 14
||||| |:
Db 5968 PTYPQRRYWDAPTG 5984

RESULT 6
US-09-335-409-2

; Sequence 2, Application US/09335409
; Patent No. 6121029
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/335,409
; CURRENT FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1421
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
; US-09-335-409-2

Query Match 47.3%; Score 43.5; DB 3; Length 1421;
Best Local Similarity 37.5%; Pred. No. 61;
Matches 9; Conservative 2; Mismatches 2; Indels 11; Gaps 1;

Qy 4 PTYPWR-----DAETGER 16
||||| |:
Db 878 PTYPQRRYWDTKADDAARGDR 901

RESULT 7
US-08-663-566A-2
; Sequence 2, Application US/08663566A
; Patent No. 5853733
; GENERAL INFORMATION:
; APPLICANT: Cochran, Mark D
; APPLICANT: Macdonald, Richard D
; TITLE OF INVENTION: Recombinant Herpesvirus of Turkeys
; TITLE OF INVENTION: and Uses Thereof
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/663,566A
; FILING DATE: June 13, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)278-0400
; TELEFAX: (212)391-0526
; TELEX: 422523
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 797 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-663-566A-2

us-09-518-931-2_copy_31_46.ra1

Wed Jan 31 13:44:28 2001

Query Match 44.6%; Score 41; DB 2; Length 797;
Best Local Similarity 43.8%; Pred. No. 81;
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 AETPTVPRDAETGER 16
| | | | | : | : |
Db 26 ASIPETPWRSTLSQOR 41

RESULT 8
US-08-023-610-2
; Sequence 2, Application US/08023610
; Patent No. 5928648
; GENERAL INFORMATION:
; APPLICANT: Cochran Ph.D., Mark D
; TITLE OF INVENTION: Recombinant Herpesvirus of Turkeys
; TITLE OF INVENTION: and Uses Thereof
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/023,610
FILING DATE: February 26, 1993

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White Esq, John P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)977-9550
TELEFAX: (212)664-0525
TELEX: 422523
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 797 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-023-610-2

Query Match 44.6%; Score 41; DB 2; Length 797;
Best Local Similarity 43.8%; Pred. No. 81;
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 AETPTVPRDAETGER 16
| | | | | : | : |
Db 26 ASIPETPWRSTLSQOR 41

RESULT 9
US-08-288-065A-2
; Sequence 2, Application US/08288065A
; Patent No. 5961982
; GENERAL INFORMATION:
; APPLICANT: Cochran, Mark D
; TITLE OF INVENTION: Recombinant Herpesvirus of Turkeys S-
; TITLE OF INVENTION: HVT-050 and Uses Thereof
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 1185 Avenue of the Americas
; CITY: New York

STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/288,065A
FILING DATE: Aug-09-94
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)278-0400
TELEFAX: (212)391-0526
TELEX: 422523
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 797 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-288-065A-2

Query Match 44.6%; Score 41; DB 2; Length 797;
Best Local Similarity 43.8%; Pred. No. 81;
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 AETPTVPRDAETGER 16
| | | | | : | : |
Db 26 ASIPETPWRSTLSQOR 41

RESULT 10
US-08-362-240A-2
; Sequence 2, Application US/08362240A
; Patent No. 5965138
; GENERAL INFORMATION:
; APPLICANT: Cochran, Mark D
; APPLICANT: Junker, David
; APPLICANT: Wild, Martha A
; TITLE OF INVENTION: Recombinant Herpesvirus and Uses Thereof
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362,240A
FILING DATE: Dec-22-94
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)278-0400
TELEFAX: (212)391-0526
TELEX: 422523
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 797 amino acids
TYPE: amino acid

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-362-240A-2

Query Match 44.6%; Score 41; DB 2; Length 797;
Best Local Similarity 43.8%; Pred. No. 81;
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 1 AETPTWPRDAETGER 16
| | | | | : : :
Db 26 ASIPETWPRSTLSGQR 41

RESULT 11

PCT-US95-10245-2
; Sequence 2, Application PC/TUS9510245
; GENERAL INFORMATION:
; APPLICANT: SYNTRO CORPORATION
; TITLE OF INVENTION: Recombinant Herpesvirus of Turkeys And Uses Thereof
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10245
; FILING DATE: 09-AUG-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)278-0400
; TELEFAX: (212)391-0526
; TELEX: 422523
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 797 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-10245-2

Query Match 44.6%; Score 41; DB 4; Length 797;
Best Local Similarity 43.8%; Pred. No. 81;
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 1 AETPTWPRDAETGER 16
| | | | | : : :
Db 26 ASIPETWPRSTLSGQR 41

RESULT 12

US-08-440-103-29
; Sequence 29, Application US/08440103
; Patent No. 5670152
; GENERAL INFORMATION:
; APPLICANT: Weiner, Amy J.
; APPLICANT: Houghton, Michael
; TITLE OF INVENTION: Immunoreactive Polypeptide Compositions
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation

; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,103
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,368
; FILING DATE:
; APPLICATION NUMBER: US 07/759,575
; FILING DATE: 13-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCRET NUMBER: 0205.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 288 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 3
; OTHER INFORMATION: /label= heterogeneity
; OTHER INFORMATION: /note= "Amino acid #3 can also be Arg."
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 7
; OTHER INFORMATION: /label= Heterogeneity
; OTHER INFORMATION: /note= "Amino Acid #5 can also be Ala."
US-08-440-103-29

Query Match 43.5%; Score 40; DB 1; Length 268;
Best Local Similarity 70.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 PTYPWRDAET 13
| | | | | : : :
Db 142 PTYNWGDNET 151

RESULT 13

US-08-440-103-30
; Sequence 30, Application US/08440103
; Patent No. 5670152
; GENERAL INFORMATION:
; APPLICANT: Weiner, Amy J.
; APPLICANT: Houghton, Michael
; TITLE OF INVENTION: Immunoreactive Polypeptide Compositions
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/440,103
  FILING DATE: 12-MAY-1995
  CLASSIFICATION: 424
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US/08/231,368
  FILING DATE:
  APPLICATION NUMBER: US 07/759,575
  FILING DATE: 13-SEP-1991
ATTORNEY/AGENT INFORMATION:
  NAME: McClung, Barbara G.
  REGISTRATION NUMBER: 33,113
  REFERENCE/DOCKET NUMBER: 0205.001
TELEPHONE: (510) 601-2708
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 30:
  LENGTH: 268 amino acids
  TYPE: amino acid
  STRANDEDNESS: single
  MOLECULE TYPE: protein
  FEATURE:
    NAME/KEY: Duplication
    LOCATION: 5
    OTHER INFORMATION: /label= Heterogeneity
    OTHER INFORMATION: /note= "This amino acid can also be Met."
  FEATURE:
    NAME/KEY: Duplication
    LOCATION: 79
    OTHER INFORMATION: /label= Heterogeneity
    OTHER INFORMATION: /note= "This amino acid can also be Val."
  FEATURE:
    NAME/KEY: Duplication
    LOCATION: 80
    OTHER INFORMATION: /label= Heterogeneity
    OTHER INFORMATION: /note= "This amino acid can also be Gly."
  FEATURE:
    NAME/KEY: Duplication
    LOCATION: 93
    OTHER INFORMATION: /label= Heterogeneity
    OTHER INFORMATION: /note= "This amino acid can also be Gln."
  FEATURE:
    NAME/KEY: Duplication
    LOCATION: 139
    OTHER INFORMATION: /label= Heterogeneity
    OTHER INFORMATION: /note= "This amino acid can only be Phe."
  FEATURE:
    NAME/KEY: Duplication
    LOCATION: 141
    OTHER INFORMATION: /label= Heterogeneity
    OTHER INFORMATION: /note= "This amino acid can also be Val."
  FEATURE:
    NAME/KEY: Duplication
    LOCATION: 191
    OTHER INFORMATION: /label= Heterogeneity
    OTHER INFORMATION: /note= "This amino acid can also be Ala."
  FEATURE:
    NAME/KEY: Duplication
    LOCATION: 197
    OTHER INFORMATION: /label= Heterogeneity
    OTHER INFORMATION: /note= "This amino acid can also be Thr."
  FEATURE:
    NAME/KEY: Duplication
    LOCATION: 208
    OTHER INFORMATION: /label= Heterogeneity
    OTHER INFORMATION: /note= "This amino acid can also be Arg and Asp."
  FEATURE:
    NAME/KEY: Duplication

```

; OTHER INFORMATION: /note= "Amino acid #3 can also be Arg."
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 7
; OTHER INFORMATION: /label= Heterogeneity
; OTHER INFORMATION: /note= "Amino Acid #5 can also be Ala."
US-08-440-542-29

Query Match 43.5%; Score 40; DB 1; Length 268;
Best Local Similarity 70.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 PTYWRDAET 13
||| | | |
Db 142 PTYKNGDNET 151

RESULT 15
US-08-440-542-30
; Sequence 30, Application US/08440542
; Patent No. 5670153
; GENERAL INFORMATION:
; APPLICANT: Weiner, Amy J.
; APPLICANT: Houghton, Michael
; TITLE OF INVENTION: Immunoreactive Polypeptide Compositions
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,542
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,368
; FILING DATE:
; APPLICATION NUMBER: US 07/759,575
; FILING DATE: 13-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0205.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 268 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 5
; OTHER INFORMATION: /label= Heterogeneity
; OTHER INFORMATION: /note= "This amino acid can also be Met."
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 79
; OTHER INFORMATION: /label= Heterogeneity
; OTHER INFORMATION: /note= "This amino acid can also be Val."
; FEATURE:

; NAME/KEY: Duplication
; LOCATION: 80
; OTHER INFORMATION: /label= Heterogeneity
; OTHER INFORMATION: /note= "This amino acid can also be Gly."
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 93
; OTHER INFORMATION: /label= Heterogeneity
; OTHER INFORMATION: /note= "This amino acid can also be Gln."
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 139
; OTHER INFORMATION: /label= Heterogeneity
; OTHER INFORMATION: /note= "This amino acid can only be Phe."
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 141
; OTHER INFORMATION: /label= Heterogeneity
; OTHER INFORMATION: /note= "This amino acid can also be Val."
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 191
; OTHER INFORMATION: /label= Heterogeneity
; OTHER INFORMATION: /note= "This amino acid can also be Ala."
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 197
; OTHER INFORMATION: /label= Heterogeneity
; OTHER INFORMATION: /note= "This amino acid can also be Thr."
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 208
; OTHER INFORMATION: /label= Heterogeneity
; OTHER INFORMATION: /note= "This amino acid can also be Arg and Asp."
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 233
; OTHER INFORMATION: /label= Heterogeneity
; OTHER INFORMATION: /note= "This amino acid can also be Trp."
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 247
; OTHER INFORMATION: /label= Heterogeneity
; OTHER INFORMATION: /note= "This amino acid can also be Lys."
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 251
; OTHER INFORMATION: /label= Heterogeneity
; OTHER INFORMATION: /note= "This amino acid can also be Gly."
US-08-440-542-30

Query Match 43.5%; Score 40; DB 1; Length 268;
Best Local Similarity 70.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 PTYWRDAET 13
||| | | |
Db 142 PTYKNGDNET 151

Search completed: January 30, 2001, 16:47:41
Job time: 617 sec

us-09-518-931-2_copy_31_46.rai

Wed Jan 31 13:44:28 2001

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:45:28 ; Search time 183.67 seconds
(without alignments)
24.574 Million cell updates/sec

Title: us-09-518-931-4_copy_1_132

Perfect score: 759

Sequence: 1 MRALEGPGLSLLCLVLPALP.....CRCTGFFAHAGFCLEHASC 132

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq_36:*
- 1: /SIDSI/gcgdata/geneseq/geneseq/AA1980.DAT:*
 - 2: /SIDSI/gcgdata/geneseq/geneseq/AA1981.DAT:*
 - 3: /SIDSI/gcgdata/geneseq/geneseq/AA1982.DAT:*
 - 4: /SIDSI/gcgdata/geneseq/geneseq/AA1983.DAT:*
 - 5: /SIDSI/gcgdata/geneseq/geneseq/AA1984.DAT:*
 - 6: /SIDSI/gcgdata/geneseq/geneseq/AA1985.DAT:*
 - 7: /SIDSI/gcgdata/geneseq/geneseq/AA1986.DAT:*
 - 8: /SIDSI/gcgdata/geneseq/geneseq/AA1987.DAT:*
 - 9: /SIDSI/gcgdata/geneseq/geneseq/AA1988.DAT:*
 - 10: /SIDSI/gcgdata/geneseq/geneseq/AA1989.DAT:*
 - 11: /SIDSI/gcgdata/geneseq/geneseq/AA1990.DAT:*
 - 12: /SIDSI/gcgdata/geneseq/geneseq/AA1991.DAT:*
 - 13: /SIDSI/gcgdata/geneseq/geneseq/AA1992.DAT:*
 - 14: /SIDSI/gcgdata/geneseq/geneseq/AA1993.DAT:*
 - 15: /SIDSI/gcgdata/geneseq/geneseq/AA1994.DAT:*
 - 16: /SIDSI/gcgdata/geneseq/geneseq/AA1995.DAT:*
 - 17: /SIDSI/gcgdata/geneseq/geneseq/AA1996.DAT:*
 - 18: /SIDSI/gcgdata/geneseq/geneseq/AA1997.DAT:*
 - 19: /SIDSI/gcgdata/geneseq/geneseq/AA1998.DAT:*
 - 20: /SIDSI/gcgdata/geneseq/geneseq/AA1999.DAT:*
 - 21: /SIDSI/gcgdata/geneseq/geneseq/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	759	100.0	170	19 W63623	Human tumour necro
2	759	100.0	245	20 Y28449	A human tumour nec
3	759	100.0	300	19 W66102	Amino acid sequenc
4	759	100.0	300	19 W63622	Human tumour necro
5	759	100.0	300	20 Y03099	Human lung TNF-rec
6	759	100.0	300	20 Y42182	Human FLINT #1 pro
7	759	100.0	300	20 Y17479	Mammalian tumour n
8	759	100.0	300	20 Y06817	Human DCR3 polypep
9	759	100.0	300	20 W97749	Human tumour necro
10	759	100.0	300	20 W95082	Orphan receptor (H
11	748	98.6	302	20 Y42183	Human FLINT #2 pro
12	745	98.2	300	21 Y77458	Human TNF receptor

13	616	81.2	271	20	Y42184	Human mFLINT #1 pr
14	605	79.7	273	20	Y42185	Human mFLINT #2 pr
15	282	37.2	153	20	Y22222	Human TNFR superfa
16	278	36.6	215	20	W93585	Human HPO6 protei
17	264	34.8	349	20	W83928	Human FTHMA-070 pa
18	262.5	34.6	401	17	R99933	Mutated OCIF, OCIF
19	261.5	34.5	143	17	R99946	Mutated OCIF, OCIF
20	261.5	34.5	187	17	R99950	Mutated OCIF, OCIF
21	261.5	34.5	197	17	R99945	Mutated OCIF, OCIF
22	261.5	34.5	272	17	R99944	Mutated OCIF, OCIF
23	261.5	34.5	321	17	R99949	Mutated OCIF, OCIF
24	261.5	34.5	326	17	R99940	Mutated OCIF, OCIF
25	261.5	34.5	327	17	R99941	Mutated OCIF, OCIF
26	261.5	34.5	351	17	R99943	Mutated OCIF, OCIF
27	261.5	34.5	359	17	R99939	Mutated OCIF, OCIF
28	261.5	34.5	393	17	R99948	Mutated OCIF, OCIF
29	261.5	34.5	399	17	R99942	Mutated OCIF, OCIF
30	261.5	34.5	401	17	R99932	Mutated OCIF, OCIF
31	261.5	34.5	401	17	R99934	Mutated OCIF, OCIF
32	261.5	34.5	401	17	R99925	Full length osteoc
33	261.5	34.5	401	17	R99931	Mutated OCIF, OCIF
34	261.5	34.5	401	19	W53239	Human OCIF genome
35	260.5	34.3	145	17	R99930	Osteoclastogenesis
36	260.5	34.3	187	21	Y77464	Primate protein se
37	260.5	34.3	390	17	R99357	Human tumour necro
38	260.5	34.3	395	19	W57636	Modified TRI recep
39	260.5	34.3	401	18	W38345	Human osteoprotege
40	260.5	34.3	401	19	W57635	TRI receptor prote
41	260.5	34.3	401	20	Y05742	Tumour necrosis fa
42	260.5	34.3	401	20	W95030	Tumour necrosis fa
43	260.5	34.3	401	20	W83926	Human FTHMA-070 pr
44	260.5	34.3	401	21	Y43400	Osteoprotegrin pro
45	258.5	34.1	154	17	R99929	Osteoclastogenesis

ALIGNMENTS

RESULT	1
W63623	
ID	W63623 standard; Protein; 170 AA.
XX	
AC	W63623;
XX	
DT	26-OCT-1998 (first entry)
XX	
DE	Human tumour necrosis factor receptor-6 beta protein.
XX	
KW	Human tumour necrosis factor receptor-6 alpha; TNFR-6 alpha; TNFR-6 beta; endothelial cells; keratinocytes; normal prostate; apoptosis;
KW	prostate tumour tissue.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Peptide
FT	Protein
FT	Region
FT	note= "TNFR-6 beta"
FT	/note= "Soluble extracellular domain"
XX	
PN	W09830694-A2.
XX	
PD	16-JUL-1998.
XX	
PF	13-JAN-1998; 98WO-US00153.
XX	
PR	14-JAN-1997; 97US-0035496.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Ebner R, Feng P, Gentz RL, Ni J, Ruben SM, Yu G;
XX	

CC regulation and to regulate the maturation of TNF ligand-bearing
CC cells such as T- or B-cells, lymphocytes, peripheral blood
CC mononuclear cells, polymorphonuclear leukocytes, fibroblasts or
CC haematopoietic cells.
XX
SQ Sequence 300 AA;

Query Match 100.0%; Score 759; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.2e-58; Indels 0; Gaps 0;
Matches 132; Conservative 0; Mismatches 0;

QY 1 MRALEGPGLSLLCLVLPALLPVPVAVRGVAETPTYPWRDAETGERLVCAQCPPGTFVQR 60
Db 1 mralegpglsllclvlpallpvpavrgvaetptypwrdaetgerlvcaqcpptfvqr 60
QY 61 PCRRDSPPTCGPCPPRHVYTFWNYLERCRYCNVLCGEREERACHATHNRCRCRTGFF 120
Db 61 pcrdspttcgpcpprhytqfwnylercrycnvlgereearachathnrcrcrtgff 120
QY 121 AHAGFCLEHASC 132
Db 121 ahagfclehasc 132

RESULT 10
W95082
ID W95082 standard; Protein; 300 AA.
XX
AC W95082;
XX
DT 20-MAY-1999 (first entry)
XX
DE Orphan receptor (HUMAN NTR-1) polypeptide.
XX
KW HUMAN NTR-1; orphan receptor; osteoprotegerin; OPB; TNFR; human;
KW tumour necrosis factor receptor; muscle disorder; bone mass; screening;
KW muscle metabolism; binding agent; cognate ligand.
XX
OS Homo sapiens.
XX
PN WO9907738-A2.
XX
PD 18-FEB-1999.
XX
PF 04-AUG-1998; 98WO-US16202.
XX
PR 06-AUG-1997; 97US-0054869.
XX
PA (PROC) PROCTER & GAMBLE CO.
PA (REGE-) REGENERON PHARM INC.
XX
PI Masiakowski PJ, Morris J, Valenzuela DM;
XX
XX WPI: 1999-167365/14.
DR N-PSDB; X22300.
XX
PT Novel orphan human receptor polypeptide and nucleic acid - useful as
PT diagnostic reagents and for treatment of muscle disorders
XX
PS Claim 7; Page 21; 23pp; English.
XX
CC This represents a HUMAN NTR-1 polypeptide, a novel orphan receptor. The
CC protein is related to osteoprotegerin (OPB) and to tumour necrosis factor
CC receptor (TNFR). Host cells transfected with a vector comprising the
CC HUMAN NTR-1 nucleic acid are used for the recombinant expression of the
CC protein. HUMAN NTR-1 proteins and antibodies immuno specific for the
CC protein are useful for diagnosis and treatment of humans and animals,
CC especially muscle disorders; as the receptor is involved in regulation of
CC bone mass and muscle metabolism. HUMAN NTR-1 receptors are also useful
CC for screening for novel binding agents, and cognate ligands, which may be
CC used to treat disorders associated with HUMAN NTR-1 imbalance.

Db 61 pcrdspttcgpcpprhytqfwnylercrycnvlgereearachathnrcrcrtgff 120
QY 121 AHAGFCLEHASC 132
Db 121 ahagfclehasc 132

RESULT 9
W97749
ID W97749 standard; Protein; 300 AA.
XX
AC W97749;
XX
DT 21-MAY-1999 (first entry)
XX
DE Human tumour necrosis factor receptor ZTNFR-5.
XX
KW ZTNFR-5; tumour necrosis factor receptor; TNFR; human;
KW cell maturation; bone cell regulation.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..23 /note= "signal peptide"
FT Protein 24..300
FT Domain 24..194 /note= "mature protein"
FT Region 49..71 /note= "extracellular domain"
FT Region 72..113 /note= "cysteine-rich pseudo-repeat 1"
FT Region 114..151 /note= "cysteine-rich pseudo-repeat 1"
FT Region 152..194 /note= "cysteine-rich pseudo-repeat 1"
FT Region /note= "cysteine-rich pseudo-repeat 1"
XX
PN WO9904001-A1.
XX
XX 28-JAN-1999.
XX
XX 21-JUL-1998; 98WO-US15072.
XX
XX 21-JUL-1997; 97US-0053203.
XX
XX (ZYMO) ZYMOGENETICS INC.
XX
XX Farrah TM;
XX
XX WPI: 1999-132245/11.
DR N-PSDB; X07226.
XX
PT Novel tumour necrosis factor receptor ZTNFR5 - useful for
PT regulating maturation of TNF-ligand bearing cells
XX
XX Claim 1; Page 84-85; 109pp; English.
XX
CC This polypeptide comprises a new, secreted tumour necrosis factor
CC receptor (see W97749), designated ZTNFR-5. Novel ZTNFR-5 encoding
CC polynucleotides and polypeptides were initially identified by
CC querying an expressed sequence tag (EST) database for sequences
CC homologous to conserved motifs within the TNF receptor family.
CC Based on this search, a contig of 16 ESTs (see X07226) was
CC constructed. ZTNFR-5 polypeptides comprise 4 cysteine-rich repeats
CC (see also W9750-55) that are homologous to other TNF receptors, in
CC particular the soluble, secreted TNF receptor osteoprotegerin.
CC ZTNFR-5 polypeptide can be prepared by recombinant methods. The
CC polypeptide, especially the extracellular domain, can be used to
CC generate a soluble variant of ZTNFR-5. The polypeptides and
CC nucleic acids can be used to screen for ligands, agonists and
CC antagonists of ZTNFR-5. The polypeptides can be used in bone cell

SQ Sequence 300 AA;

Query Match 100.0%; Score 759; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 2.2e-58; Indels 0; Gaps 0;
 Matches 132; Conservative 0; Mismatches 0;

QY 1 MRALEGGLSLCLVLPALLPVPVAVRGVAETPTYPWRDAETGERLVCAQCPTGTFVQR 60
 |||||
 DB 1 mraleppglsilclvlpallpvpavrgvaetpypwrdaetgerlvcaqcpgtfvqr 60
 |||||

QY 61 PCRRDPTTCGPPRHYTQFWNYLRCRYCNVLCGEREEARACHATHNRCRTGFF 120
 |||||
 DB 61 pccrdspttcgpprhytqfwnylercrycnvlgereearachathnrcrcrtgff 120
 |||||

QY 121 AHAGFGLCHASC 132
 |||||
 DB 121 ahagfglehasc 132

RESULT 11
 Y42183
 ID Y42183 standard; Protein; 302 AA.
 XX
 AC Y42183;
 XX
 DT 17-DEC-1999 (first entry)
 XX
 DE Human FLINT #2 protein sequence.
 XX
 KW Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
 KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
 KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
 KW reperfusion-associated injury; aplastic anaemia; differentiation;
 KW growth; myelodysplastic syndrome; pancytopenic condition;
 KW myocardial ischaemia.
 KW
 XX Homo sapiens.
 XX
 OS Homo sapiens.
 XX
 PN W09950413-A2.
 XX
 PD 07-OCT-1999.
 XX
 PF 30-MAR-1999; 99WO-US06797.
 XX
 PR 30-MAR-1998; 98US-0079856.
 PR 20-MAY-1998; 98US-0086074.
 PR 09-SEP-1998; 98US-0099643.
 PR 17-DEC-1998; 98US-0112577.
 PR 18-DEC-1998; 98US-0112703.
 PR 18-DEC-1998; 98US-0112933.
 PR 22-DEC-1998; 98US-0113407.
 XX
 PA (ELIL) LILLY & CO ELI.
 XX
 PI Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
 PI Hui KY, Khaitonenkov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
 PI Song HY, Wang J, Wu X, Zuckerman SH;
 XX
 DR WPI; 1999-591319/50.
 DR N-PSDB; 225376.
 XX
 PT Use of mature FLINT for treating acute liver failure, inflammation,
 PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
 PT and proinflammatory activity
 XX
 XX Example 2; Fig 2; 99pp; English.
 XX
 CC The present invention describes therapeutic applications of mature FLINT
 CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
 CC (mFLINT), which is a member of the tumour necrosis factor receptor
 CC superfamily, is used for treating acute liver failure, inflammation of
 CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated

CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
 CC injury or disorder such as hypercoagulation (including use with
 CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
 CC or disorder. Type I diabetes, cancer, cell damage or damage to an
 CC innocent bystander tissue that is induced by a chemotherapeutic agent or
 CC therapeutic irradiation, treating haematopoietic progenitor cells that
 CC have been exposed to therapeutic radiation or chemotherapy, aplastic
 CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
 CC also used for promoting the growth or differentiation of a haematopoietic
 CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
 CC resulting from abnormal myocardial ischaemia. The present sequence
 CC represents human FLINT.
 XX
 SQ Sequence 302 AA;

Query Match 98.6%; Score 748; DB 20; Length 302;
 Best Local Similarity 98.5%; Pred. No. 1.9e-57;
 Matches 132; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 1 MRALEGGLSLCLVLPALLPVPVAVRGVAETPTYPWRDAETGERLVCAQCPTGTFVQR 60
 |||||
 DB 1 mraleppglsilclvlpallpvpavrgvaetpypwrdaetgerlvcaqcpgtfvqr 60
 |||||

QY 61 PCRRDPTTCGPPRHYTQFWNYLRCRYCNVLCGEREEARACHATHNRA--CRCRTG 118
 |||||
 DB 61 pccrdspttcgpprhytqfwnylercrycnvlgereearachathnrcrcrtg 120
 |||||

QY 119 FFAHAGFGLCHASC 132
 |||||
 DB 121 ffahagfglehasc 134

RESULT 12
 Y77458
 ID Y77458 standard; Protein; 300 AA.
 XX
 AC Y77458;
 XX
 DT 05-JUN-2000 (first entry)
 XX
 DE Human TNF receptor-like protein, HDTEA84.
 XX
 KW TNF receptor family; tumour necrosis factor; HDTEA84; HSLJD37R;
 KW Rank-like protein; RANKL; immune disorder; inflammation; allergy;
 KW immunosuppressant; antiarthritic; antirheumatoid; antiinflammatory;
 KW dermatological; antithyroid.
 XX
 OS Homo sapiens.
 XX
 PN W0200001817-A2.
 XX
 PD 13-JAN-2000.
 XX
 PF 06-JUL-1999; 99WO-US12366.
 XX
 PR 06-JUL-1998; 98US-0110938.
 PR 13-JUL-1998; 98US-0114466.
 PR 23-JUL-1998; 98US-0093897.
 PR 12-AUG-1998; 98US-0132968.
 PR 18-AUG-1998; 98US-0136214.
 PR 11-SEP-1998; 98US-0099999.
 XX
 PA (SCHE) SCHERING CORP.
 XX
 PI Bates EEM, Lebecque SJE, Murphy EE, Mattson JD, Gorman DM;
 PI Hedrick JA, Wang L, Zlotnik A, Murgolo NJ, Greene JR, Johnston JA;
 PI Bazan JF, Mahony D, Lees EM;
 XX
 DR WPI; 2000-171015/15.
 DR N-PSDB; 292404.
 XX
 PT New isolated mammalian genes, used to develop products for treating

e.g., immune, inflammatory or allergic abnormalities, cancers or degenerative conditions

Claim 24; Page 157; 218pp; English.

The invention relates to a number of primate and/or rodent proteins, and the genes which encode them. The invention encompasses human dendritic cell prostaglandin transporter (DC-PGT); the TNF (tumour necrosis factor) receptor family-related proteins HDTEA84, HSLJD37R and RANKL; human CC chemokine HCC5; human deubiquitinating proteins Dub11 and Dub 12; human MD-1 and human and murine MD-2 proteins, which exhibit the properties of ligands for proteins comprising a leucine-rich motif (LRR); human cyclin E2; cDNAs encoding these proteins; and antibodies against these proteins. The proteins can be used for modulating the physiology or development of a cell. They can be used to mediate uptake of substrates (e.g., prostaglandin-like molecules), to modulate or mediate cellular interactions (e.g., induce or prevent trafficking, proliferation, or differentiation of cells), or are intracellular proteins which are important in various cellular processes such as the deubiquitination of proteins or cell cycle regulation. The products can be used for treating medical conditions such as immune, inflammatory or allergic disorders, or abnormal cellular proliferation, for example, cancers or degenerative conditions. They can be used to modulate immune responses in disease states e.g., autoimmune disorders, including rheumatoid arthritis, systemic lupus erythematosus, Hashimoto's autoimmune thyroiditis, as well as acute and chronic inflammatory responses in which T cell activation, expansion, and/or immunological T cell memory play an important role. Sequences Y77458-Y77461 and Y77465-Y77468 represent TNF receptor family-related proteins. Y77458 is the human protein HDTEA84, Y77459-Y77461 are human HSLJD37R proteins, Y77465 is murine Rank-like protein RANKL, and Y77466-Y77468 are human RANKL proteins.

Sequence 300 AA;

Query Match 98.2%; Score 745; DB 21; Length 300;
Best Local Similarity 98.5%; Pred. No. 3.5e-57;
Matches 130; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRALGGGLSLLCLVLALLPVPVAVRGVAETPTYPWRDAETGERLVCAQCPCPFVQR 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1 mralepgglsllclvlallpvpvavrgvaetptypwrdaetgerlvcaqcpqtfvqr 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 61 PCRRDSPTCGPPRHYYTQFWNYLERCYCNVLCGEREEERACHATHNRCACRTGFF 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 61 pcrrdspmtcpcpprhytqfwnylercyvcylcgereearachathnrcacrtgff 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 121 AHAGFCLEHASC 132
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 121 ahagfclehasc 132
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 13
Y42184
ID Y42184 standard; Protein: 271 AA.
XX XX
AC AC
XX XX
DE DE
DT DT
XX XX
DE DE
XX XX
KW KW
KW KW
KW KW
KW KW
XX XX
OS OS
PN PN

Human mFLINT #1 protein sequence.
Human; FLINT; OPG3; tumour necrosis factor receptor; FasL; apoptosis; inflammation; cancer; diabetes; acute liver failure; sepsis; hepatitis; ischaemia-associated injury; hypercoagulation; reperfusion-associated injury; aplastic anaemia; differentiation; growth; myelodysplastic syndrome; pancytopenic condition; myocardial ischaemia.
Homo sapiens.
W09950413-A2.

07-OCT-1999.
30-MAR-1999; 99WO-US06797.
30-MAR-1998; 98US-0079856.
20-MAY-1998; 98US-0086074.
09-SEP-1998; 98US-0099643.
17-DEC-1998; 98US-0112577.
18-DEC-1998; 98US-0112703.
22-DEC-1998; 98US-0112933.
22-DEC-1998; 98US-0113407.
(ELJIL) LILLY & CO ELI.
Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
Hui KY, Kharitonov A, Mirzahi J, Na S, Noblitt TW, Reidy CA;
Song HY, Wang J, Wu X, Zuckerman SH;
WPI; 1999-591319/50.
N-PSDB; Z25377.
Use of mature FLINT for treating acute liver failure, inflammation,
cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
and proinflammatory activity
Claim 31; Fig 3; 99pp; English.
The present invention describes therapeutic applications of mature FLINT
(mFLINT) for use in the treatment of acute liver failure. Mature FLINT
(mFLINT), which is a member of the tumour necrosis factor receptor
superfamily, is used for treating acute liver failure, inflammation of
the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
injury or disorder such as hypercoagulation (including use with
thrombolytic or anti-thrombotic agents), reperfusion-associated injury
or disorder, Type I diabetes, cancer, cell damage or damage to an
innocent bystander tissue that is induced by a chemotherapeutic agent or
therapeutic irradiation, treating haematopoietic progenitor cells that
have been exposed to therapeutic radiation or chemotherapy, aplastic
anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
also used for promoting the growth or differentiation of a haematopoietic
progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
resulting from abnormal myocardial ischaemia. The present sequence
represents human mFLINT.

Sequence 271 AA;

Query Match 81.2%; Score 616; DB 20; Length 271;
Best Local Similarity 100.0%; Pred. No. 4.2e-46;
Matches 103; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 VAETPTYWRDAETGERLVCAQCPCPGTFVQRPCRRDSPTCGCPPRHYTQFWNYLERCR 89
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1 vaetptywrdaetgerlvcaqcpqgtfvqrpcrrdspsttcgcpcprhytqfwnyer 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 90 YCNVLCGEREEERACHATHNRCACRTGFFAHAGFCLEHASC 132
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 61 ycnvlgereearachathnrcacrtgffahagfclehasc 103
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 14
Y42185
ID Y42185 standard; Protein: 273 AA.
XX XX
AC AC
XX XX
DT DT
XX XX
DE DE
XX XX
KW KW

Human mFLINT #2 protein sequence.
Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;

KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
KW reperfusion-associated injury; aplastic anaemia; differentiation;
KW growth; myelodysplastic syndrome; pancytopenic condition;
KW myocardial ischaemia.
XX
OS Homo sapiens.
XX
PN W09950413-A2.
XX
PD 07-OCT-1999.
XX
XX 30-MAR-1999; 99WO-US06797.
XX
XX 30-MAR-1998; 98US-0079856.
PR 20-MAY-1998; 98US-0086074.
PR 09-SEP-1998; 98US-0099643.
PR 17-DEC-1998; 98US-0112577.
PR 18-DEC-1998; 98US-0112703.
PR 18-DEC-1998; 98US-0112933.
PR 22-DEC-1998; 98US-0113407.
XX
XX (ELIL) LILLY & CO ELI.
XX
XX Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
PI Song HY, Wang J, Wu X, Zuckerman SH;
XX
DR WPI; 1999-591319/50.
DR N-PSDB; 225378.
XX
XX Use of mature FLINT for treating acute liver failure, inflammation,
PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
PT and proinflammatory activity
XX
XX Example 2; Fig 4; 99pp; English.
XX
XX The present invention describes therapeutic applications of mature FLINT
CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
CC (mFLINT), which is a member of the tumour necrosis factor receptor
CC superfamily, is used for treating acute liver failure, inflammation of
CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
CC injury or disorder such as hypercoagulation (including use with
CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
CC or disorder, Type I diabetes, cancer, cell damage or damage to an
CC innocent bystander tissue that is induced by a chemotherapeutic agent or
CC therapeutic irradiation, treating haematopoietic progenitor cells that
CC have been exposed to therapeutic radiation or chemotherapy, aplastic
CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
CC also used for promoting the growth or differentiation of a haematopoietic
CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
CC resulting from abnormal myocardial ischaemia. The present sequence
XX represents human mFLINT.
XX
SQ Sequence 273 AA;

Query Match 79.7%; Score 605; DB 20; Length 273;
Best Local Similarity 98.1%; Pred. No. 3.8e-45;
Matches 103; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

Qy 30 VAETPTYPWRDAETGERLVCAQCPGTGVQRCRDSPTTCGCPPRHYTQFWNYLRCR 89
|||||
Db 1 vaetptyprdaetgerlvcaqcpgtgvqrcrdspttcgcpccprhytqfwnylerc 60
|||||

Qy 90 YCNVLCGEREREACHATHNRA--CRCTGFFAHAGFCLHASC 132
|||||
Db 61 ycnvlgereereachathnracrcrtctgffahagfclhasc 105
|||||

RESULT 15
Y22222

ID Y22222 standard; Protein; 153 AA.
XX
AC Y22222;
XX
DT 16-SEP-1999 (first entry)
XX
DE Human TNFR superfamily soluble receptor protein sequence.
XX
XX TNFL1; human; TNFR superfamily; tumour necrosis factor ligand; TNF;
KW tumour necrosis factor receptor; TNFR superfamily; cell proliferation;
KW cell differentiation; cytokine production; immunoglobulin; hyperplasia;
KW apoptosis inducer; activated T cell; autoimmune disease; inhibitor;
KW myasthenia gravis; insulin-dependent diabetes mellitus; endotoxic shock;
KW rheumatoid arthritis; multiple sclerosis; systemic lupus erythematosus;
KW tumour; proliferative disorder; neoplasia; dysplasia; immunocompetence;
KW lymphoid organogenesis; bacterial resistance; contact hypersensitivity;
KW delayed type sensitivity; therapy.
OS Homo sapiens.
XX
XX W09933980-A2.
PN
XX 08-JUL-1999.
PD
XX 22-DEC-1998; 98WO-US27474.
XX
XX 16-DEC-1998; 98US-0212270.
PR 30-DEC-1997; 97US-0068959.
XX
XX (CHIR) CHIRON CORP.
XX
XX Kassam A, Lamson G, Pot D, Tribouley C;
XX WPI; 1999-405508/34.
DR N-PSDB; X84621.
XX
XX New tumour necrosis factor ligands, useful for induction of cell
PT death and/or proliferation of cells
PT
XX Claim 1; Page 61; 69pp; English.
XX
XX This sequence represents a tumour necrosis factor receptor (TNFR)
CC superfamily soluble protein of the invention. The invention also relates
CC to tumour necrosis factor (TNF) ligand (TNFL) family proteins. The TNFL
CC proteins play regulatory roles in cell proliferation and/or
CC differentiation, e.g. they can induce production of cytokines,
CC immunoglobulins, etc. A variety of diseases can be treated by modulating
CC the activity of TNFL proteins, e.g. they can induce apoptosis of
CC activated T cells but rescue resting T cell from apoptosis. TNFL
CC polypeptides can therefore be used to treat autoimmune diseases, such as
CC myasthenia gravis, insulin-dependent diabetes mellitus, rheumatoid
CC arthritis, multiple sclerosis, and systemic lupus erythematosus. TNFL
CC proteins also have tumour stimulating properties, so tumours can be
CC treated by inhibiting the expression or activity of TNFL. Other
CC proliferative disorders, such as neoplasias, dysplasias, and hyperplasia
CC can also be treated using TNFL inhibitors. The TNFL polypeptides and
CC polynucleotides can also be used to enhance or decrease TNF activity,
CC thus providing therapeutic benefits such as induction of cell death,
CC lymphoid organogenesis, or host bacterial resistance, and inhibition of
CC endotoxic shock, contact hypersensitivity, delayed type sensitivity or
CC immunocompetence of a transplant recipient. TNF and its receptors play a
CC major role in host defence and immunosurveillance. As such, there is a
CC need to identify new members of TNFR families. This invention provides
CC this need.
XX
SQ Sequence 153 AA;

Query Match 37.2%; Score 282; DB 20; Length 153;
Best Local Similarity 100.0%; Pred. No. 1.6e-17;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 85 LERCRCNVLCGEREREACHATHNRCRCRTGFFAHAGFCLHASC 132

Db 1 lercrycnvlgereeeearachathnracrctgffahagfclchasc 48

Search completed: January 30, 2001, 16:45:29
Job time: 589 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:50:33 ; Search time 149.64 Seconds
(without alignments)
59.896 Million cell updates/sec

Title: US-09-518-931-4_COPY_1_132

Perfect score: 759

Sequence: 1 MRALGPGLSLLCLVLALPA.....CRCRTGFFAHAGFCLEHASC 132

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries.

Database :

PIR_66:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	191.5	25.2	435	2 I54182	tumor necrosis fac
2	189	24.9	459	2 I48854	gene murine tumour
3	186	24.5	474	2 B38634	tumor necrosis fac
4	185	24.4	349	2 D72175	G2R protein - vari
5	185	24.4	349	2 D36858	gene G4R protein -
6	184.5	24.3	348	2 T28623	hypothetical prote
7	176	23.2	461	1 A35356	tumor necrosis fac
8	152	20.0	326	1 GOV2ML	T2 protein - myxom
9	151.5	20.0	325	2 B43692	T2 protein - rabbi
10	141.5	18.6	277	2 I37552	OX40 homolog - hum
11	140	18.4	260	1 A46517	CD27 antigen precu
12	125.5	16.5	324	2 JC2395	Fas antigen precu
13	124	16.3	327	2 A46484	apoptosis-mediati
14	122.5	16.1	454	1 GOMST1	tumor necrosis fac
15	122.5	16.1	454	2 I57826	tumor necrosis fac
16	121.5	16.0	271	2 S12783	OX40 antigen precu
17	118	15.5	272	2 I48700	gene ox40 protein
18	118	15.5	314	2 I37383	FAS soluble protei
19	118	15.5	335	2 A40036	apoptosis-mediati
20	116	15.3	255	2 I38426	4-1BB - human
21	116	15.3	255	2 JT0752	lymphocyte activat
22	114	15.0	250	1 A49053	CD27 antigen precu
23	113.5	15.0	256	2 B32393	T-cell antigen 4-1
24	111.5	14.7	416	1 JN0006	nerve growth facto
25	109	14.4	595	2 A42086	CD30 antigen precu
26	108.5	14.3	455	1 GQHTT1	tumor necrosis fac
27	104.5	13.8	461	1 GQHTT1	tumor necrosis fac
28	103.5	13.6	425	1 A26431	nerve growth facto
29	102.5	13.5	5376	2 T42215	zonadhesin - mouse

30	101.5	13.4	461	2 JC4302	tumor necrosis fac
31	96.5	12.7	427	1 GQHUN	nerve growth facto
32	96	12.6	277	2 A60771	B-cell activation
33	92.5	12.2	1299	2 T43251	furin (EC 3.4.21.7
34	91	12.0	289	2 A46515	B cell-associated
35	91	12.0	305	2 A46476	B cell-associated
36	89	11.7	293	2 T22919	hypothetical prote
37	89	11.7	2321	2 S78549	notch3 protein - h
38	88.5	11.7	493	2 JC5486	membrane glycoprot
39	88	11.6	1019	2 T13039	tyrosine kinase re
40	88	11.6	1548	2 S34583	serine proteinase
41	88	11.6	1620	2 T27283	hypothetical prote
42	87.5	11.5	1574	2 T13954	MEGF6 protein - ra
43	87.5	11.5	2610	2 T20968	hypothetical prote
44	87	11.5	2476	2 T34022	zonadhesin - pig
45	86.5	11.4	895	2 T49010	hypothetical prote

ALIGNMENTS

RESULT 1

I54182

tumor necrosis factor receptor 2-related protein - human

C:Species: Homo sapiens (man)

C>Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 17-Mar-2000

C:Accession: I54182

R:Baens, M.; Chafranet, M.; Cassiman, J.J.; Van den Berghe, H.; Marynen, P.

Genomics 16, 214-218, 1993

A:Title: Construction and evaluation of a hncDNA library of human l2p transcribed seq

A:Reference number: I54182; MUID:93252381

A:Accession: I54182

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-435 <RES>

A:Cross-references: GB:L04270; NID:g339761; PIDN:AAA36757.1; PID:g339762

C:Genetics:

A:Gene: GDB:LTBR

A:Cross-references: GDB:1230195; OMIM:600979

A:Map position: l2p13.3-l2p13.1

C:Superfamily: tumor necrosis factor receptor type 1; NGF receptor repeat homology

Query Match	25.2%	Score 191.5;	DB 2;	Length 435;
Best Local Similarity	35.0%;	Pred. No. 1.8e-09;		
Matches	50;	Conservative	13;	Mismatches 61;
				Indels 19;
				Gaps 6;
Qy	3	ALEGPGLSLLCLVLALPALPVPVAVRGVAETPTV-----PWRDA-----ETGERLVCAQC	52	
Db	6	ATSAPGLAWPLVLGLFGLLAASQPAV---PPYASENQTCRDQEKYEYEPQHRICCSRC	62	
Qy	53	PPGTFVQRCRRDSPTTCGPPRHYTFQFNYL---ERCRCNVLCGREGREBEARACHATH	109	
Db	63	PPGTYYSAKCSRIKRDTCATCAENSNEHWNLTQLCRPCDPVMG--LEEIAPTSKR	120	
Qy	110	NRACRCRTGFFAHAGFCLEHASC	132	
Db	121	KTQCRCPGMFC-AAWALECTHC	142	

RESULT 2

I48854

gene murine tumour necrosis factor receptor 2 protein - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 23-Jul-1999

C:Accession: I48854

R:Powell, E.E.; Wicker, L.S.; Peterson, L.B.; Todd, J.A.

Mamm. Genome 5, 726-727, 1994

A:Title: Allelic variation of the type 2 tumor necrosis factor receptor gene.

A:Reference number: I48854; MUID:95178848

A:Accession: I48854

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

C:Genetics:

A:Gene: G4R

C:Superfamily: myxoma virus T2 protein; NGF receptor repeat homology

F:32-66/Domain: NGF receptor repeat homology <NGF>

F:68-109/Domain: NGF receptor repeat homology <NG2>

F:110-151/Domain: NGF receptor repeat homology <NG3>

Query Match 24.4%; Score 185; DB 2; Length 349;

Best Local Similarity 29.2%; Pred. No. 5.4e-09;

Matches 38; Conservative 15; Mismatches 65; Indels 12; Gaps 2;

Qy 9 LSLLCLVLPALPVPVAVRGVAETPTYPWRDAE-TGERLVCAQCPCPGTFVQPCRRDSTP 68

Db 10 LFLSCLIIINGRDAAPYTPNGKCKDTEY-----KRHNLCCLSCPPGYASRLCDSKINT 63

Qy 69 TCGPCPPRHRYTFQWNYLERCRYCNVLCGEREEERACHATHNACRCHRTGFF-----AH 122

Db 64 OCTPCGSGTFTSRNNHLPACLSCNGRCNSQVETRSCNTTHNRICECPGYCYLLKGSSG 123

Qy 123 AGFCLEHASC 132

Db 124 CKACVSQTKC 133

RESULT 6

T28623

hypothetical protein G2R - variola major virus

C:Species: variola major virus

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 21-Jul-2000

C:Accession: T28623

R:Massung, R.F.; Esposito, J.J.; Liu, L.I.; Qi, J.; Utterback, T.R.; Knight, J.C.; Aubin

Nature 366, 748-751, 1993

A:Title: Potential virulence determinants in terminal regions of variola smallpox virus

A:Reference number: Z20488; MUID:94088747

A:Accession: T28623

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-348 <MAS>

A:Cross-references: EMBL:L22579; NID:g623595; PIDN:AAA0933.1; PID:g439102

A:Experimental source: strain Bangladesh 1975

C:Superfamily: myxoma virus T2 protein; NGF receptor repeat homology

Query Match

Best Local Similarity 30.5%; Score 184.5; DB 2; Length 348;

Matches 40; Conservative 14; Mismatches 62; Indels 15; Gaps 3;

Qy 9 LSLLCLVLPALPVPVAVRGVAETPTYPWRDAE-TGERLVCAQCPCPGTFVQPCRRDSTP 67

Db 10 LFLSCLIIINGRDAAPYTP-----PNGKCKDTEYKRHNLCCLSCPPGYASRLCDSKTN 61

Qy 68 TTGCGPCPPRHRYTFQWNYLERCRYCNVLCGEREEERACHATHNACRCHRTGFF-----A 121

Db 62 TQCTPCGSGTFTSRNNHLPACLSCNGRCNSQVETRSCNTTHNRICECPGYCYLLKGSS 121

Qy 122 HAGFCLEHASC 132

Db 122 GCKACVSQTKC 132

RESULT 7

A35356

tumor necrosis factor receptor type 2 precursor - human

N:Alternate names: 75K tumor necrosis factor receptor

C:Species: Homo sapiens (man)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: A35356; A36475; A48416; A36007; A23666; B35010; I38094

R:Smith, C.A.; Davis, T.; Anderson, D.; Solam, L.; Beckmann, M.P.; Jerzy, R.; Dower, S.K

Science 248, 1019-1023, 1990

A:Title: A receptor for tumor necrosis factor defines an unusual family of cellular and

A:Reference number: A35356; MUID:90260639

A:Accession: A35356

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-461 <SMI>

A:Cross-references: GB:M32315; NID:g189185; PIDN:AAA59929.1; PID:g189186

R:Kohn, T.; Brewer, M.T.; Baker, S.L.; Schwartz, P.E.; King, M.W.; Hale, K.K.; Squir

Proc. Natl. Acad. Sci. U.S.A. 87, 8331-8335, 1990

A:Title: A second tumor necrosis factor receptor gene product can shed a naturally oc

A:Reference number: A36475; MUID:91045991

A:Accession: A36475

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-195,'R',197-461 <KOH>

A:Cross-references: GB:M55994; GB:M38549; NID:g339757; PIDN:AAA36755.1; PID:g339758

R:Dembic, Z.; Loetscher, H.; Gubler, U.; Pan, Y.C.; Lahm, H.W.; Gentz, R.; Brockhaus,

Cytokine 2, 231-237, 1990

A:Title: Two human TNF receptors have similar extracellular, but distinct intracellular

A:Reference number: A48416; MUID:91370690

A:Accession: A48416

A:Status: preliminary

A:Molecule type: mRNA; protein

A:Residues: 23-461 <DEM>

A:Cross-references: GB:S63368; NID:g235648; PIDN:AA19824.1; PID:g235649

A:Note: sequence extracted from NCBI backbone (NCBIN:63368, NCBIIP:63371)

R:Heller, R.A.; Song, K.; Onasch, M.A.; Fischer, W.H.; Chang, D.; Ringold, G.M.

Proc. Natl. Acad. Sci. U.S.A. 87, 6151-6155, 1990

A:Title: Complementary DNA cloning of a receptor for tumor necrosis factor and demons

A:Reference number: A36007; MUID:90349572

A:Accession: A36007

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 116-140,'P',142-195,'R',197-362,'T',364-461 <HEL>

A:Cross-references: GB:M35857; NID:g339751; PIDN:AAA63262.1; PID:g339752

R:Loetscher, H.; Schlaeger, E.J.; Lahm, H.W.; Pan, Y.C.E.; Lesslauer, W.; Brockhaus,

J. Biol. Chem. 265, 20131-20138, 1990

A:Title: Purification and partial amino acid sequence analysis of two distinct tumor

A:Reference number: A23666; MUID:91056048

A:Accession: A23666

A:Status: preliminary

A:Molecule type: protein

A:Residues: 23-40;65-89;136-141;300-306 <LOE>

R:Engelmann, H.; Novick, D.; Wallach, D.

J. Biol. Chem. 265, 1531-1536, 1990

A:Title: Two tumor necrosis factor-binding proteins purified from human urine. Eviden

A:Reference number: A35010; MUID:90110215

A:Accession: B35010

A:Status: preliminary

A:Molecule type: protein

A:Residues: 27-31 <ENG>

R:Kuhnert, P.; Kemper, O.; Wallach, D.

Gene 150, 381-386, 1994

A:Title: Cloning, sequencing and partial functional characterization of the 5' region

A:Reference number: I38094; MUID:95121934

A:Accession: I38094

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-37 <RES>

A:Cross-references: EMBL:X80021; NID:g666044; PIDN:CAA56324.1; PID:g825701

C:Genetics:

A:Gene: GDB:TNFR2

A:Cross-references: GDB:I25914; OMIM:191191

A:Map position: lp36.2-1p36.2

A:Introns: 26/3

A:Note: the list of introns is incomplete

C:Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology

C:Keywords: duplication; glycoprotein; receptor; transmembrane protein

F:1-22/Domain: signal sequence #status predicted <SIG>

F:23-416/Product: tumor necrosis factor receptor type 2 #status experimental <MAT>

F:40-76/Domain: NGF receptor repeat homology <NG1>

F:78-119/Domain: NGF receptor repeat homology <NG2>

F:120-162/Domain: NGF receptor repeat homology <NG3>

F:164-201/Domain: NGF receptor repeat homology <NG4>

F:262-279/Domain: transmembrane #status predicted <TMN>

F:280-461/Domain: intracellular #status predicted <INT>

A:Cross-references: GB:M63928; NID:g180084; PIDN:AAA58411.1; PID:g180085
A:Note: sequence extracted from NCBI backbone (NCBIN:60285, NCBIP:60289)
C:Comment: A soluble CD27 found in serum and urine is formed by proteolysis.
C:Genetics:

A:Gene: GDB:CD27
A:Cross-references: GDB:I32582; OMIM:186711
A:Map position: 12p13-12p13
A:Introns: 46/1; 90/1; 150/1; 180/1; 220/1
C:Superfamily: CD27 antigen; NGF receptor repeat homology
C:Keywords: duplication; glycoprotein; homodimer; phosphoprotein; receptor; surface anti
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-260/Product: CD27 antigen #status predicted <EXT>
F:21-191/Domain: extracellular #status predicted <EXT>
F:27-63/Domain: NGF receptor repeat homology <NG1>
F:65-105/Domain: NGF receptor repeat homology <NG2>
F:121-188/Region: proline/serine/threonine-rich
F:192-211/Domain: transmembrane #status predicted <TMN>
F:212-280/Domain: intracellular #status predicted <INT>
F:95/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 18.4%; Score 140; DB 1; Length 260;
Best Local Similarity 31.0%; Pred. No. 3.3e-05;
Matches 39; Conservative 14; Mismatches 49; Indels 24; Gaps 7;

Oy 12 LCLVLALLPVPVAVRGVAETPTYPWRDAETGERLVCAQCPPGTFVORPCRRDPTT-C 70

Db 9 LCVGLTGLSLATPAKSCPE--RHYWAQG----KLCQMCPEPTFLVKDCDQHRKTAQC 62

Oy 71 GPC-----PPRHYTFQWNYLERCRYN--VLGEREEEARACHATHNRACRRTGFFAH 122

Db 63 DPCIPGVFSFDHTR--PHCESCRHNSGLL-----VRNCTITANAECACRNGWQCR 113

Oy 123 AGFCLE 128

Db 114 DRECTE 119

RESULT 12

JC2395
Fas antigen precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 05-Nov-1999
C:Accession: JC2395; PC2246
R:Kimura, K.; Wakatsuki, T.; Yamamoto, M.
Biochem. Biophys. Res. Commun. 198, 666-674, 1994
A:Title: A variant mRNA species encoding a truncated form of Fas antigen in the rat live
A:Reference number: JC2395; MUID:94128114
A:Accession: JC2395
A:Molecule type: mRNA
A:Residues: 1-324 <KIM>
A:Cross-references: DDBJ:D26112; NID:g468486; PIDN:BAA05108.1; PID:d1005650; PID:g468487
A:Experimental source: thymus
A:Accession: PC2246
A:Molecule type: mRNA
A:Residues: 1-62, 'RFT' <K12>
A:Cross-references: DDBJ:D26113; NID:g468488; PIDN:BAA05109.1; PID:d1005651; PID:g468489
C:Genetics:

A:Introns: 62/1
C:Superfamily: NGF receptor repeat homology
C:Keywords: transmembrane protein
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-324/Product: Fas antigen #status predicted <MAT>
F:44-79/Domain: NGF receptor repeat homology <NGF>
F:81-124/Domain: NGF receptor repeat homology <NG4>
F:171-188/Domain: transmembrane #status predicted <TMN>

Query Match 16.5%; Score 125.5; DB 2; Length 324;
Best Local Similarity 27.5%; Pred. No. 0.00069;
Matches 42; Conservative 13; Mismatches 59; Indels 39; Gaps 9;

Oy 11 LCLVLALLPVPVAV-----RGVAETPTYPWRDAETGERL----- 47
Db 1 MLWIMAVPLVLVLAVGELNVRMQGTDIFEGLELKKSVRET-----DNNCSGLYQVGPF 54

Oy 48 VCAQCPPGTFVORPCRRD-SPTTCGCPPP-RHYTFQWNYLERCRYCNVLCGERE--EAR 103

Db 55 CCQPCQCPGERVKDCCTTSGGAPTCPCTEGEYTDKKHYSDKCRC-AFCDEGHGLEVET 113

Oy 104 ACHATHNRACRRTGFFAHAGFCLEH----ASC 132

Db 114 NCTRTQNTKCKCKENFYCNASLC-DHCYHCTSC 145

RESULT 13

A46484

apoptosis-mediating membrane-associated polypeptide Fas - mouse

C:Species: Mus musculus (house mouse)

C:Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999

C:Accession: A46484; A47254

R:Watanabe-Fukunaga, R.; Brannan, C.I.; Itoh, N.; Yonehara, S.; Copeland, N.G.; Jenki

J. Immunol. 148, 1274-1279, 1992

A:Title: The cDNA structure, expression, and chromosomal assignment of the mouse Fas

A:Reference number: A46484; MUID:92148151

A:Accession: A46484

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-327 <WAT>

A:Cross-references: GB:M83649; NID:g193225; PIDN:AAA37593.1; PID:g193226

A:Experimental source: BAM3 macrophage cell line

A:Note: sequence extracted from NCBI backbone (NCBIN:81544, NCBIP:81545)

R:Adachi, M.; Watanabe-Fukunaga, R.; Nagata, S.

Proc. Natl. Acad. Sci. U.S.A. 90, 1756-1760, 1993

A:Title: Aberrant transcription caused by the insertion of an early transposable elem

A:Reference number: A47254; MUID:93189576

A:Accession: A47254

A>Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-96 <ADA>

A:Cross-references: GB:S56490; NID:g298505; PIDN:AAB25700.1; PID:g298506

A:Experimental source: MRL lpr/lpr

A:Note: sequence extracted from NCBI backbone (NCBIN:126850, NCBIN:126853, NCBIN:1268

C:Superfamily: NGF receptor repeat homology

C:Keywords: transmembrane protein

F:44-79/Domain: NGF receptor repeat homology <NGF>

F:81-124/Domain: NGF receptor repeat homology <NG4>

Query Match 16.3%; Score 124; DB 2; Length 327;
Best Local Similarity 33.0%; Pred. No. 0.00094;
Matches 32; Conservative 10; Mismatches 45; Indels 10; Gaps 6;

Oy 44 GERLVCAQCPPGTFVORPCRRDPTTTCGCPPP-RHYTFQWNYLERCRYCNVLCGERE-- 99

Db 51 GGPFCCQCPGKKKVEDCKMNGGTPTCAPCTEGEYTDKKHAYADKCRRC-TLCEEHGL 109

Oy 100 BEARACHATHNRACRRTGFFAHAGFCLEH----ASC 132

Db 110 EVETNCTLTQNTKCKCKPDYCDSPGC-EHCVCASC 145

RESULT 14

GQMST1

N:Alternate names: tumor necrosis factor precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 22-Jun-1999

C:Accession: A38634; B40254; S16677; S19021; I54532

R:Lewis, M.; Tartaglia, L.A.; Lee, A.; Bennett, G.L.; Rice, G.C.; Wong, G.H.W.; Chen,

Proc. Natl. Acad. Sci. U.S.A. 88, 2830-2834, 1991

A:Title: Cloning and expression of cDNAs for two distinct murine tumor necrosis facto

A:Reference number: A38634; MUID:91187885

A:Accession: A38634

A:Molecule type: mRNA

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 17:06:08 ; Search time 80.04 seconds
(without alignments)
53.259 Million cell updates/sec

Title: US-09-518-931-4_COPY_1_132
Perfect score: 759
Sequence: 1 MRALGPGLSLCLLALPA.....CRCTGFFAHAGCFLHASC 132

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	191.5	25.2	435	1 TNRC_HUMAN	P36941 homo sapien
2	186	24.5	474	1 TNRC_MOUSE	P25119 mus musculus
3	185	24.4	349	1 TNRC2_VARV	P34015 variola vir
4	176	23.2	461	1 TNRC2_HUMAN	P20333 homo sapien
5	168	22.1	415	1 TNRC_MOUSE	P50284 mus musculus
6	152	20.0	326	1 VT2_MXVL	P29825 myxoma viru
7	151.5	20.0	325	1 VT2_SFVRA	P25943 Shope fibro
8	141.5	18.6	277	1 OX40_HUMAN	P43489 homo sapien
9	136	17.9	260	1 CD27_HUMAN	P26842 homo sapien
10	134.5	17.7	471	1 TNR1_BOVIN	O19131 bos taurus
11	131	17.3	332	1 FASA_PIG	O77736 sus scrofa
12	125.5	16.5	324	1 FASA_RAT	Q63199 rattus norv
13	125	16.5	323	1 FASA_BOVIN	P31867 bos taurus
14	124	16.3	327	1 FASA_MOUSE	P25446 mus musculus
15	122.5	16.1	454	1 TNR1_MOUSE	P25118 mus musculus
16	121.5	16.0	271	1 OX40_RAT	P15725 rattus norv
17	118	15.5	272	1 OX40_MOUSE	P47741 mus musculus
18	118	15.5	335	1 FASA_HUMAN	P25445 homo sapien
19	116	15.3	255	1 41BB_HUMAN	Q07011 homo sapien
20	114	15.0	250	1 CD27_MOUSE	P41272 mus musculus
21	113.5	15.0	256	1 41BB_MOUSE	P20334 mus musculus
22	111.5	14.7	416	1 NGFR_CHICK	P18519 gallus gall
23	109	14.4	595	1 CD30_HUMAN	P28908 homo sapien
24	108.5	14.3	455	1 TNR1_HUMAN	P19438 homo sapien
25	104.5	13.8	461	1 TNR1_RAT	P22934 rattus norv
26	103.5	13.6	425	1 NGFR_RAT	P07174 rattus norv
27	101.5	13.4	461	1 TNR1_PIG	P50555 sus scrofa
28	97	12.8	269	1 CD40_BOVIN	Q28203 bos taurus
29	96.5	12.7	427	1 NGFR_HUMAN	P08138 homo sapien
30	96	12.6	277	1 CD40_HUMAN	P25942 homo sapien
31	91	12.0	289	1 CD40_MOUSE	P27512 mus musculus
32	89.5	11.8	417	1 WSL1_HUMAN	Q93038 h wsl-1 pro
33	87	11.5	1173	1 TSPL_XENLA	P35448 xenopus lae

34 87 11.5 2476 1 ZAN_PIG Q28983 sus scrofa
35 86 11.3 103 1 VA53_VACCC P21071 vaccinia vi
36 86 11.3 103 1 VA53_VACCV P24756 vaccinia vi
37 84.5 11.1 915 1 PAC6_RAT P41413 rattus norv
38 82.5 10.9 1680 1 FUR2_DROME P30432 drosophila
39 82 10.8 727 1 NETA_VACCC Q24567 drosophila
40 81.5 10.7 122 1 VC22_VACCC P21106 vaccinia vi
41 81.5 10.7 969 1 PAC4_HUMAN P29122 homo sapien
42 81 10.7 1798 1 LMB2_HUMAN P52268 homo sapien
43 80.5 10.6 915 1 PAC6_MOUSE Q04592 mus musculus
44 80.5 10.6 1522 1 BAI3_HUMAN O60242 homo sapien
45 79.5 10.5 937 1 PAC4_RAT Q63415 rattus norv

ALIGNMENTS

RESULT 1
TNRC_HUMAN
ID TNRC_HUMAN STANDARD; PRT; 435 AA.
AC P36941;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE LYMPHOTOXIN-BETA RECEPTOR PRECURSOR (TUMOR NECROSIS FACTOR RECEPTOR
DE 2 RELATED PROTEIN) (TUMOR NECROSIS FACTOR C RECEPTOR).
GN LTBR OR TNFR OR TNFRSF3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=LIVER.
RC MEDLINE=93252381; PubMed=8486360;
RA Baens M., Chaifanet M., Cassiman J.J., den Berghe H., Marynen P.;
RT "Construction and evaluation of a hncDNA library of human l2p
RL transcribed sequences derived from a somatic cell hybrid.";
RL Genomics 16:214-218(1993).
[2]
RN FUNCTION.
RP FUNCTION.
RA MEDLINE=94225209; PubMed=8171323;
RA Crowe P.D., van Arsdale T.L., Walter B.N., Ware C.F., Hession C.,
RA Ehrenfels B., Browning J.L., Din W.S., Goodwin R.G., Smith C.A.;
RA "A lymphotoxin-beta-specific receptor.";
RL Science 264:707-710(1994).
CC -!- FUNCTION: RECEPTOR FOR THE LYMPHOTOXIN-BETA. POSSIBLE FUNCTION IN
CC IMMUNE DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; L04270; AAA36757.1; -
DR HSSP; P25942; ICDF.
DR MIN; 600979; -
DR INTERPRO; IPR001368; -
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00500; TNFR_NGFR_2; 3.
KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 30
FT CHAIN 31 435 LYMPHOTOXIN-BETA RECEPTOR.
FT DOMAIN 31 227 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 228 248 POTENTIAL.
FT DOMAIN 249 435 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 42 211 4 X TNFR-CYS.
FT REPEAT 42 81 TNFR-CYS 1.

```
FT REPEAT 82 124 TNFR-CYS 2.
FT REPEAT 125 168 TNFR-CYS 3.
FT REPEAT 169 211 TNFR-CYS 4.
FT DISULFID 43 58 BY SIMILARITY.
FT DISULFID 59 72 BY SIMILARITY.
FT DISULFID 62 80 BY SIMILARITY.
FT DISULFID 83 98 BY SIMILARITY.
FT DISULFID 101 116 BY SIMILARITY.
FT DISULFID 104 124 BY SIMILARITY.
FT DISULFID 126 132 BY SIMILARITY.
FT DISULFID 139 148 BY SIMILARITY.
FT DISULFID 142 167 BY SIMILARITY.
FT DISULFID 170 185 BY SIMILARITY.
FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 435 AA; 46709 MW; 62462656022F656F CRC64;

Query Match 25.2%; Score 191.5; DB 1; Length 435;
Best Local Similarity 35.0%; Pred. No. 1.9e-10;
Matches 50; Conservative 13; Mismatches 61; Indels 19; Gaps 6;

QY 3 ALSEGLSLICLVLPALLPVPARGVETPTTY-----PWRDA-----ETGERLVCAQC 52
Db 6 ATSAAPGLANGPLVLGLFLAASQPAV---PPYASENQTCDQKEKEYEYEPQHRICCSRC 62

QY 53 PCCTEVQRCRRDPTTCGPPRHVYTFQWNYL---ERCYCNVLCGEREEARACHATH 109
Db 63 PPGTYVSACSRRTDVCATCAENSTNEHWNLYLTICQLCRPCDPVWG--LEETAPCTSKR 120

QY 110 NRACRCRTGFFAHAGFCLEHASC 132
Db 121 KTCRCQPGMFC-AWALECTHC 142

RESULT 2
TNR2_MOUSE STANDARD; PRT; 474 AA.
AC P25119; P97893;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR 2 PRECURSOR (TNF-R2) (P75).
GN TNFRSF1B OR TNFR2 OR TNFR-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RX SEQUENCE FROM N.A.
RA MEDLINE=91187885; PubMed=1849278;
RA Lewis M., Tartaglia L.A., Lee A., Bennett G.L., Rice G.C.,
RA Wong G.H., Chen E.Y., Goeddel D.V.;
RA "Cloning and expression of cDNAs for two distinct murine tumor
RT necrosis factor receptors demonstrate one receptor is species
RT specific.;"
RL Proc. Natl. Acad. Sci. U.S.A. 88:2830-2834(1991).
RN [2]
RX SEQUENCE FROM N.A.
RA MEDLINE=91246168; PubMed=1645445;
RA Goodwin R.G., Anderson D., Jerzy R., Davis T., Brannan C.I.,
RA Copeland N.G., Jenkins N.A., Smith C.A.;
RA "Molecular cloning and expression of the type 1 and type 2 murine
RT receptors for tumor necrosis factor.;"
RL Mol. Cell. Biol. 11:3020-3026(1991).
RN [3]
RX SEQUENCE OF 1-26 FROM N.A.
RC STRAIN=NOD;
RA Jacob C.O., Liu J.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RX SEQUENCE OF 1-22 FROM N.A.
RC TISSUE=LIVER;
RA Kissoneghis M., Fellows R., Feldmann M., Chernajovsky Y.;
```

```
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: RECEPTOR FOR TNF-ALPHA.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M60469; AAA39752.1; -
DR EMBL; M59378; AAA40463.1; -
DR EMBL; U39488; AAA85021.1; -
DR EMBL; X87128; CRA60618.1; -
DR PIR; B38634; B38634.
DR HSSP; P19438; INCF.
DR MGD; MGI:1314883; TNFRSF1B.
DR INTERPRO; IPR001368; -.
DR PAM; PF00020; TNFR-C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00505; TNFR_NGFR_2; 3.
KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 22
FT CHAIN 23 474 TUMOR NECROSIS FACTOR RECEPTOR 2.
FT DOMAIN 23 258 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 259 288 POTENTIAL.
FT DOMAIN 289 474 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 39 203 4 X TNFR-CYS.
FT REPEAT 39 77 TNFR-CYS 1.
FT REPEAT 78 119 TNFR-CYS 2.
FT REPEAT 120 164 TNFR-CYS 3.
FT REPEAT 165 203 TNFR-CYS 4.
FT DISULFID 40 54 BY SIMILARITY.
FT DISULFID 55 68 BY SIMILARITY.
FT DISULFID 58 76 BY SIMILARITY.
FT DISULFID 79 94 BY SIMILARITY.
FT DISULFID 97 111 BY SIMILARITY.
FT DISULFID 101 119 BY SIMILARITY.
FT DISULFID 121 127 BY SIMILARITY.
FT DISULFID 136 145 BY SIMILARITY.
FT DISULFID 139 163 BY SIMILARITY.
FT DISULFID 166 181 BY SIMILARITY.
FT CARBOHYD 69 69 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 474 AA; 50319 MW; 462AE398C4D6563 CRC64;

Query Match 24.5%; Score 186; DB 1; Length 474;
Best Local Similarity 37.6%; Pred. No. 6.4e-10;
Matches 32; Conservative 10; Mismatches 39; Indels 4; Gaps 1;

QY 46 RLVCACQPGCTFVQRCRRDPTTCGPPRHVYTFQWNYLERCYCNVLCGEREEARAC 105
Db 52 QMCCAKCPPGQYVYKFCNKTSDTVCADCEASMTQWNNFRCLSCSSCTTDDQVEIRAC 111

QY 106 HATHNRACRCRTGFF----AHAGFC 126
Db 112 TKQNRVCAACEAGRYCALKTHGSC 136

RESULT 3
VC22_VARV STANDARD; PRT; 349 AA.
AC P34015;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE PROTEIN C22/B28 HOMOLOG.
GN G4R.
OS Variola virus.
```

```
OC Viruses: dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-INDIA-1967 / ISOLATE IND3;
RX MEDLINE-93202281; PubMed-8384129;
RA Shchelkunov S.N., Blinov V.M., Sandakhchiev L.S.;
RT "Genes of variola and vaccinia viruses necessary to overcome the host
RT protective mechanisms.";
RL FEBS Lett. 319:80-83(1993).
CC -!- SIMILARITY: CONTAINS 2 LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGIONS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X69198; CAA49137.1; -.
DR EMBL: X67117; CAA47540.1; -.
DR PIR: D36858; D36858.
DR PIR: S35987; S35987.
DR PIR: S46888; S46888.
DR HSSP: P19438; INCF.
DR INTERPRO: IPR001368; -.
DR PFAM: PF00020; TNFR_C6; 2.
DR PROSITE: PS00652; TNFR_NGFR_1; 2.
DR PROSITE: PS00050; TNFR_NGFR_2; 2.
KW Repeat.
FT DOMAIN 31 108 2 X TNFR-CYS.
FT REPEAT 31 66 TNFR-CYS 1.
FT REPEAT 67 108 TNFR-CYS 2.
SQ SEQUENCE 349 AA; 38189 MW; D45D40B5C6E780EF CRC64;

      24.4%: Score 185; DB 1; Length 349;
      Best Local Similarity 29.2%; Pred No. 6.1e-10;
      Matches 38; Conservative 15; Mismatches 65; Indels 12; Gaps 2;

QY 9 LSLCLVLALPALLPVAVRGVAETPTYPWRDAETGERLVCAQCPTGTVQRPCCRDSPT 68
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
10 LFLSCIIINGRDAAPYPNPGKCKDTEY-----KRHNLCCLSCPPTGVASRLCDSKNT 63
QY 69 TCGPCPPRHYYTFWNYLERCRVCNVLGEREEAEARACHATHNRACRTGFF-----AH 122
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
64 QCTPCGSGTFTSRNNHLPACLSGRCNCSNQVETRSCNTHNRICECSPGYCYLLKGSSG 123
QY 123 AGFCLEHASC 132
Db | | | | |
124 CRACVSQTKC 133

RESULT 4
TNFR2_HUMAN
ID TNFR2_HUMAN STANDARD; PRT; 461 AA.
AC P20333;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR 2 PRECURSOR (TUMOR NECROSIS FACTOR
DE BINDING PROTEIN 2) (TBPII) (P80) (TNF-R2) (P75) (CD120B) (ETANERCEPT).
GN TNFRSF1B OR TNFR2 OR TNFRB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-90260639; PubMed-2160731;
RA Smith C.A., Davis T., Anderson D., Solam L., Beckmann M.P., Jerzy R.,
RA Dower S.K., Cosman D., Goodwin R.G.;
RT "A receptor for tumor necrosis factor defines an unusual family of
```

```
RT cellular and viral proteins.";
RL Science 248:1019-1023(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-91045991; PubMed-2172983;
RA Kohno T., Brewer M.T., Baker S.L., Schwartz P.E., King M.W.,
RA Hale K.K., Squires C.H., Thompson R.C., Vannice J.L.;
RT "A second tumor necrosis factor receptor gene product can shed a
RT naturally occurring tumor necrosis factor inhibitor.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:8331-8335(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE-96299745; PubMed-8661109;
RA Beltinger C.P., White P.S., Maris J.M., Sulman E.P., Jensen S.J.,
RA Lepassier D., Stallard B.J., Goeddel D.V., Desauvage F.J.,
RA Brodeur G.M.;
RT "Physical mapping and genomic structure of the human TNFR2 gene.";
RL Genomics 35:94-100(1996).
RN [4]
RP SEQUENCE OF 116-461 FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE-90349572; PubMed-2166946;
RA Heller R.A., Song K., Onasch M.A., Fischer W.H., Chang D.,
RA Ringold G.M.;
RT "Complementary DNA cloning of a receptor for tumor necrosis factor
RT and demonstration of a shed form of the receptor.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:6151-6155(1990).
RN [5]
RP SEQUENCE OF 27-31.
RX MEDLINE-90110215; PubMed-2153136;
RA Engelmann H., Novick D., Wallach D.;
RT "Two tumor necrosis factor-binding proteins purified from human
RT urine. Evidence for immunological cross-reactivity with cell surface
RT tumor necrosis factor receptors.";
RL J. Biol. Chem. 265:1531-1536(1990).
RN [6]
RP SEQUENCE OF 22-40; 65-69; 136-141; 300-306 AND 346-362.
RX MEDLINE-91056048; PubMed-2173696;
RA Loetscher H., Schlaeger E.J., Lahm H.-W., Pan Y.-C.E., Lesslauer W.,
RA Brockhaus M.;
RT "Purification and partial amino acid sequence analysis of two
RT distinct tumor necrosis factor receptors from HL60 cells.";
RL J. Biol. Chem. 265:20131-20138(1990).
RN [7]
RP CHARACTERIZATION.
RX MEDLINE-93016040; PubMed-1328224;
RA Pennica D., Lam V.T., Mize N.K., Weber R.F., Lewis M., Fendly B.M.,
RA Lipari M.T., Goeddel D.V.;
RT "Biochemical properties of the 75-kDa tumor necrosis factor receptor.
RT Characterization of ligand binding, internalization, and receptor
RT phosphorylation.";
RL J. Biol. Chem. 267:21172-21178(1992).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 419-428 IN TRAF2 COMPLEX.
RX MEDLINE-99221490; PubMed-10206649;
RA Park Y.C., Burkitt V., Villa A.R., Tong L., Wu H.;
RT "Structural basis for self-association and receptor recognition of
RT human TRAF2.";
RL Nature 398:533-538(1999).
CC -!- FUNCTION: RECEPTOR FOR TNF-ALPHA. HIGH AFFINITY FOR TNF-ALPHA AND
CC APPROXIMATELY 5-FOLD LOWER AFFINITY FOR TNF-BETA.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- PTM: PHOSPHORYLATED; MAINLY ON SERINE RESIDUES WITH A VERY LOW
CC LEVEL ON THREONINE RESIDUES.
CC -!- PHARMACEUTICAL: AVAILABLE UNDER THE NAME ENBREL (IMMUNEX AND
CC WYETH-AYERIS). USED TO TREAT MODERATE TO SEVERE RHEUMATOID
CC ARTHRITIS (RA). ENBREL CONSIST OF THE EXTRACELLULAR LIGAND-BINDING
CC PORTION OF TNFR2 LINKED TO AN IMMUGLOBULIN FC CHAIN. IT BINDS TO
CC TNF-ALPHA AND BLOCKS ITS INTERACTIONS WITH RECEPTORS.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -!- DATABASE: NAME-PROW; NOTE-CD guide cd120b entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd120b.htm".
CC -!- DATABASE: NAME-Enbrel; NOTE-Clinical information on Enbrel;
CC WWW="http://www.enbrelinfo.com/".
```



```
FT DISULFID 101 116 BY SIMILARITY.
FT DISULFID 104 124 BY SIMILARITY.
FT DISULFID 126 132 BY SIMILARITY.
FT DISULFID 139 150 BY SIMILARITY.
FT DISULFID 142 169 BY SIMILARITY.
FT DISULFID 172 187 BY SIMILARITY.
FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 415 AA; 44956 MW; 29B326A566AEF661 CRC64;

Query Match 22.1%; Score 168; DB 1; Length 415;
Best Local Similarity 36.1%; Pred. No. 2.4e-08;
Matches 44; Conservative 15; Mismatches 47; Indels 16; Gaps 6;

QY 6 GP---GLSLGLLVALPALLPVPVAVGVAETPTYPWRDAETGERL---VCAQCPPTGVQ 59
  || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 15 GPLLLGLSL-LVASQPOLVPPYRI---ENQTCWDQDKEYEYEPMDHVDCCSRCPGGEVF 69
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 60 RPCRDSPTTCGPPRHRYTQFNVL---ERCRCYCNVLCGEREEARACHATHNRACRCR 116
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 70 AVCSQDTVCKTCHPNSYNEHNHLSTCQLCRPCDIVLG--FEEVAPCTSDRAECRCQ 127
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 117 TG 118
  |
Db 128 PG 129

RESULT 6
VT2_MYXVL STANDARD; PRT; 326 AA.
AC P29825;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE TUMOR NECROSIS FACTOR SOLUBLE RECEPTOR PRECURSOR (PROTEIN T2).
GN T2.
OS Myxoma virus (strain Lausanne).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Leporipoxvirus.
[1]
RN RP SEQUENCE FROM N.A.
RX MEDLINE=91335768; PubMed=1651597;
RA Upton C., Macen J.L., Schreiber M., McFadden G.;
RT "Myxoma virus expresses a secreted protein with homology to the tumor
RT necrosis factor receptor gene family that contributes to viral
RT virulence.";
RL Virology 184:370-382(1991).
CC -!- FUNCTION: BINDS TO TNF-ALPHA AND BETA. PROBABLY PREVENTS TNF TO
CC REACH CELLULAR TARGET AND THEREBY DEAMPENING THE POTENTIAL
CC ANTIVIRAL EFFECTS OF THE CYTOKINE.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M95181; AAA46632.1; .
DR EMBL; A23729; CAA01688.1; .
DR PIR; A40566; GQVZML.
DR HSSP; P19438; 1TNR.
DR INTERPRO; IPR001368; .
DR PFAM; PF00020; TNFR_C6; 2.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00505; TNFR_NGFR_2; 2.
KW Receptor; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 326 TUMOR NECROSIS FACTOR SOLUBLE RECEPTOR.
FT DOMAIN 27 186 4 X TNFR-CYS.
```

```
FT REPEAT 27 62 TNFR-CYS 1.
FT REPEAT 63 104 TNFR-CYS 2.
FT REPEAT 105 147 TNFR-CYS 3.
FT REPEAT 148 186 TNFR-CYS 4.
FT CARBOHYD 66 66 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 326 AA; 35208 MW; ABBF027E947292FF CRC64;

Query Match 20.0%; Score 152; DB 1; Length 326;
Best Local Similarity 30.6%; Pred. No. 5.4e-07;
Matches 37; Conservative 12; Mismatches 56; Indels 16; Gaps 3;

QY 12 LCLVALPALI-----PVPVAVGVAETPTYPWRDAETGERLVCAQCPGTFVQRCRDS 66
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4 LTLLEYVACVYGGAPYGADRGKRGNDY-----EKDGLCCTSCPPGVSASRLCGPGS 57
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 67 PTTGCPCPRHRYTQFNVLRCRYCNVLCGEREEARACHATHNRACRCRTGFFAHAGFC 126
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 58 DTVCSCKNETFTASTNHAPACVSCRGCRGTGHLSESQCDKTRDRVDCSAG-----NYC 112
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 127 L 127
  |
Db 113 L 113

RESULT 7
VT2_SFVKA STANDARD; PRT; 325 AA.
AC P25943;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE TUMOR NECROSIS FACTOR SOLUBLE RECEPTOR PRECURSOR (PROTEIN T2).
GN T2.
OS Shope fibroma virus (strain Kasza) (SFV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Leporipoxvirus.
[1]
RN RP SEQUENCE FROM N.A.
RX MEDLINE=87321103; PubMed=2820128;
RA Upton C., Delange A.M., McFadden G.;
RT "Tumorigenic poxviruses: genomic organization and DNA sequence of the
RT telomeric region of the Shope fibroma virus genome.";
RL Virology 160:20-30(1987).
[2]
RN RP FUNCTION.
RP MEDLINE=91207415; PubMed=1850261;
RA Smith C.A., Davis T., Wignall J.M., Din W.S., Farrah T., Upton C.,
RA McFadden G., Goodwin R.G.;
RT "T2 open reading frame from the Shope fibroma virus encodes a soluble
RT form of the TNF receptor.";
RL Biochem. Biophys. Res. Commun. 176:335-342(1991).
CC -!- FUNCTION: BINDS TO TNF-ALPHA AND BETA. PROBABLY PREVENTS TNF TO
CC REACH CELLULAR TARGET AND THEREBY DEAMPENING THE POTENTIAL
CC ANTIVIRAL EFFECTS OF THE CYTOKINE.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M17433; .; NOT_ANNOTATED_CDS.
DR EMBL; A23727; CAA01687.1; .
DR PIR; B43692; B43692.
DR HSSP; P19438; 1TNR.
DR INTERPRO; IPR001368; .
```

[illegible]

RX MEDLINE-92195401; PubMed-1372394;
RA Watanabe-Fukunaga R., Brannan C.I., Copeland N.G., Jenkins N.A.,
RA Nagata S.;
RT "Lymphoproliferation disorder in mice explained by defects in Fas
RT antigen that mediates apoptosis.";
RL Nature 356:314-317(1992).
CC -!- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS FASL. THE
CC ADAPTOR MOLECULE FADD RECRUITS CASPASE-8 TO THE ACTIVATED
CC RECEPTOR. THE RESULTING AGGREGATE CALLED THE DEATH-INDUCING
CC SIGNALING COMPLEX (DISC) PERFORMS CASPASE-8 PROTEOLYTIC
CC ACTIVATION. ACTIVE CASPASE-8 INITIATES THE SUBSEQUENT CASCADE OF
CC CASPASES (ASPARTATE-SPECIFIC CYSTEINE PROTEASES) MEDIATING
CC APOPTOSIS. FAS-MEDIATED APOPTOSIS MAY HAVE A ROLE IN THE
CC INDUCTION OF PERIPHERAL TOLERANCE, IN THE ANTIGEN-STIMULATED
CC SUICIDE OF MATURE T-CELLS, OR BOTH (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: DETECTED IN VARIOUS TISSUES INCLUDING THYMUS,
CC LIVER, LUNG, HEART, AND ADULT OVARY.
CC -!- DOMAIN: CONTAINS A DEATH DOMAIN INVOLVED IN THE BINDING OF FADD,
CC AND MAYBE TO OTHER CYTOSOLIC ADAPTOR PROTEINS.
CC -!- DISEASE: DEFECTS IN FAS ARE THE CAUSE OF A LYMPHOPROLIFERATION
CC DISORDER (LPR) RESPONSIBLE FOR LYMPHADENOPATHY AND AUTOANTIBODY
CC PRODUCTION.
CC -!- SIMILARITY: CONTAINS A LA-NGER/TNFR-TYPE CYSTEINE-RICH REGION.
CC -!- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M83649; AAA37593.1; -;
DR EMBL; S56490; AAB25700.1; -;
DR EMBL; S56485; AAB25700.1; JOINED.
DR EMBL; S56486; AAB25700.1; JOINED.
DR PIR; A46484; A46484.
DR HSP; P25445; 1DDF.
DR MGD; MGI:95484; FAS.
DR INTERPRO; IPR000488; -;
DR INTERPRO; IPR001368; -;
DR PFAM; PF00020; TNFR_c6; 3.
DR PFAM; PF00531; Death; 1.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00652; TNFR_NGFR_2; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
KW Apoptosis; Receptor; Glycoprotein; Transmembrane; Repeat; Signal;
KW Disease mutation.
FT SIGNAL 1 21
FT CHAIN 22 327 FASL RECEPTOR.
FT DOMAIN 22 169 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 170 186 POTENTIAL.
FT DOMAIN 187 227 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 43 162 3 X TNFR-CYS.
FT REPEAT 43 79 TNFR-CYS 1.
FT REPEAT 80 123 TNFR-CYS 2.
FT REPEAT 124 162 TNFR-CYS 3.
FT DOMAIN 222 306 DEATH DOMAIN.
FT CARBOHYD 43 43 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 114 114 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARIANT 246 246 I -> N (IN LPR).
FT SEQUENCE 327 AA; 37418 MW; F6BFFC5ACE356EEE CRC64;

Query Match 16.3%; Score 124; DB 1; Length 327;
Best Local Similarity 33.0%; Pred. No. 0.00018;
Matches 32; Conservative 10; Mismatches 45; Indels 10; Gaps 6;
QY 44 GRLVCAOCPPTVQRPQRDSPT--TCGPCPP-RHYTFQWNLRCRCYNVLCGERE-- 99
DB 51 GGFPCQPCQPCQKKVEDCKMNGGPTCAPCTEGSKYMDKNHYADKCRRC-TLCDHEHCL 109

QY 100 BEARACHATHNACRCRTGFFFAHAGFCLHEH----ASC 132
DB 110 EVETNCTLTQNTKCKCKPFDYCDSPGC-EHCVCASC 145
RESULT 15
TNRI_MOUSE
ID TNRI_MOUSE STANDARD; PRT; 454 AA.
AC P25118;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR 1 PRECURSOR (p60) (TNF-R1) (TNF-RI)
DE (P55).
GN TNFRSF1A OR TNFR1 OR TNFR-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-91187885; PubMed-1849278;
RA Lewis M., Tartaglia L.A., Lee A., Bennett G.L., Rice G.C.,
RA Wong G.H., Chen E.Y., Goeddel D.V.;
RT "Cloning and expression of cDNAs for two distinct murine tumor
RT necrosis factor receptors demonstrate one receptor is species
RT specific.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:2830-2834(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-91246168; PubMed-1645445;
RA Goodwin R.G., Anderson D., Jerzy R., Davis T., Brannan C.I.,
RA Copeland N.G., Jenkins N.A., Smith C.A.;
RT "Molecular cloning and expression of the type 1 and type 2 murine
RT receptors for tumor necrosis factor.";
RL Mol. Cell. Biol. 11:3020-3026(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE-91285014; PubMed-1647956;
RA Barrett K., Taylor-Fishwick D.A., Cope A.P., Kissoneghis A.M.,
RA Gray P.W., Feldmann M., Foxwell B.M.J.;
RT "Cloning, expression and cross-linking analysis of the murine p55
RT tumor necrosis factor receptor.";
RL Eur. J. Immunol. 21:1649-1656(1991).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE-SPLEEN;
RX MEDLINE-92039815; PubMed-1657766;
RA Rothe J.G., Brockhaus M., Gentz R., Lesslauer W.;
RT "Molecular cloning and expression of the mouse Tnf receptor type b.";
RL Immunogenetics 34:338-340(1991).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE-94245292; PubMed-8188324;
RA Bebo B.F., Linthicum D.S.;
RT "Nucleotide sequence of the TNF type I receptor from a mouse
RT endothelioma cell line.";
RL Immunogenetics 39:450-451(1994).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE-93156721; PubMed-8381516;
RA Rothe J., Bluthmann H., Gentz R., Lesslauer W., Steinmetz M.;
RT "Genomic organization and promoter function of the murine tumor
RT necrosis factor receptor beta gene.";
RL Mol. Immunol. 30:165-175(1993).
CC -!- FUNCTION: RECEPTOR FOR TNF-ALPHA. THE ADAPTOR MOLECULE FADD
CC RECRUITS CASPASE-8 TO THE ACTIVATED RECEPTOR. THE RESULTING
CC AGGREGATE CALLED THE DEATH-INDUCING SIGNALING COMPLEX (DISC)
CC PERFORMS CASPASE-8 PROTEOLYTIC ACTIVATION WHICH INITIATES THE
CC SUBSEQUENT CASCADE OF CASPASES (ASPARTATE-SPECIFIC CYSTEINE
CC PROTEASES) MEDIATING APOPTOSIS (BY SIMILARITY).
CC -!- SUBUNIT: TNF BINDING TO THE EXTRACELLULAR DOMAIN OF TNFR1 LEADS TO
CC HOMOTRIMERIZATION. ONCE AGGREGATED THE RECEPTORS DEATH DOMAINS

us-09-518-931-4_copy_1_132.rsp

Wed Jan 31 13:45:04 2001

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:54:57 ; Search time 254.27 Seconds
(without alignments)
60.847 Million cell updates/sec

Title: US-09-518-931-4_COPY_1_132

Perfect score: 759

Sequence: 1 MRALGPGLSLLCLVLPALP.....CRCRTGFFAHAGFCLEHASC 132

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_15:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organellae:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	759	100.0	300	4	O95407
2	260.5	34.3	401	4	O00300
3	257	33.9	372	4	Q9UHP4
4	246	32.4	401	11	O08712
5	243	32.0	401	11	O08727
6	204.5	26.9	302	13	Q9PUS0
7	195	25.7	348	12	O57277
8	195	25.7	349	12	O57100
9	193	25.4	348	12	O57103
10	193	25.4	348	12	O57108
11	193	25.4	349	12	O57291
12	193	25.4	349	12	O57101
13	193	25.4	349	12	O57102
14	192.5	25.4	349	12	O57099
15	189	24.9	459	11	Q62327
16	185	24.4	349	12	O57284
17	185	24.4	349	12	O57098
18	185	24.4	349	12	O57110
19	185	24.4	349	12	O57111

20	185	24.4	349	12	O89118	O89118 variola vir
21	185	24.4	349	12	O89098	O89098 variola vir
22	184.5	24.3	348	12	O57112	O57112 variola vir
23	184.5	24.3	348	12	O85407	O85407 variola vir
24	182.5	24.0	482	11	O88734	O88734 mus musculus
25	181	23.8	349	12	O57097	O57097 camelpox vi
26	181	23.8	351	12	O57117	O57117 cowpox viru
27	181	23.8	355	12	O85308	O85308 cowpox viru
28	180	23.7	351	12	O73559	O73559 cowpox viru
29	179.5	23.6	326	12	O57120	O57120 cowpox viru
30	179.5	23.6	350	12	O57116	O57116 cowpox viru
31	177	23.3	349	12	O57109	O57109 variola vir
32	176	23.2	326	12	O57122	O57122 cowpox viru
33	176	23.2	349	12	O57305	O57305 cowpox viru
34	175	23.1	350	12	O57123	O57123 cowpox viru
35	173	22.8	347	12	O57119	O57119 cowpox viru
36	173	22.8	360	12	O57118	O57118 cowpox viru
37	170	22.4	347	12	O57115	O57115 cowpox viru
38	170	22.4	351	12	O57121	O57121 cowpox viru
39	168	22.1	655	4	O75509	O75509 homo sapien
40	165	21.7	439	4	Q16042	Q16042 homo sapien
41	155	20.4	322	12	O72761	O72761 cowpox viru
42	154	20.3	320	12	O57079	O57079 cowpox viru
43	153	20.2	316	12	O57092	O57092 ectromelia
44	152	20.0	320	12	O57091	O57091 ectromelia
45	152	20.0	320	12	O57300	O57300 ectromelia

ALIGNMENTS

RESULT	1
O95407	
ID	O95407
AC	O95407;
DT	01-MAY-1999 (TrEMBLrel. 10, Created)
DT	01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT	01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE	DECOY RECEPTOR 3 (M68) (M68C) (M68E).
GN	DCR3 OR TR6.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=99087326; PubMed=9872321;
RA	Pitti R.M., Marsters S.A., Lawrence D.A., Roy M., Kischkel F.C.,
RA	Dowd P., Huang A., Donahue C.J., Sherwood S.W., Baldwin D.T.,
RA	Godowski P.J., Wood W.I., Gurney A.L., Hillan K.J., Cohen R.L.,
RA	Goddard A.D., Botstein D., Ashkenazi A.
RT	"Genomic amplification of a decoy receptor for Fas ligand in lung and
RT	colon cancer.";
RL	Nature 396:699-703(1998).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Blood;
RA	MEDLINE=99253915; PubMed=10318773;
RA	Yu K.Y., Kwon B., Ni J., Zhai Y., Ebner R., Kwon B.S.;
RT	"A newly identified member of tumor necrosis factor receptor
RT	superfamily (TR6) suppresses LIGHT-mediated apoptosis.";
RL	J. Biol. Chem. 274:13733-13736(1999).
RN	[3]
RP	SEQUENCE FROM N.A.
RC	TISSUE=pancreas;
RA	MEDLINE=20122600; PubMed=10655513;
RA	Bai C., Connolly B., Metzker M.L., Hilliard C.A., Liu X., Sandig V.,
RA	Soderman A., Galloway S.M., Liu Q., Austin C.P., Caskey C.T.;
RT	"Overexpression of M68/Dcr3 in human gastrointestinal tract tumors
RT	independent of gene amplification and its location in a four-gene
RT	cluster.";
RL	Proc. Natl. Acad. Sci. U.S.A. 97:1230-1235(2000).
DR	EMBL; AF104419; AAD03056.1; -.

```

7      11 LCLVLALPALLPVAVRGAETH--PPYWRDAETGERLVCACQCPDGTVPQRCDSDPT 68
9          | | | : | | | | | | | | | | | | | | | | | | | | | | | |
        4 LLCCAL---VFLDISIKWTTQTTPPKYLHYDEETSQLLCOKCPPGTYLKQHCTAKWKI 60

7      69 TCGCPPRPHYTQFWNYLERGRYNLCGEEREEARACHATHNRACRCRTGFFAHAGFCLE 128
9          | | | : | | | : | | | : | | | : | | | : | | | : | | | :
        61 VCACPDPHYTDSNHTSDECLYCSPVKELQYKQBECNRTHNRVCECKEGRYLETEFCLK 120

7      129 HASC 132
9          | | |
        121 HRSC 124

```


Q9UHP4 PRELIMINARY; PRT; 372 AA.
AC Q9UHP4;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE OSTEOPROTEGERIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID-9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA He Z.-Y., Yang G.-Z., Zhang W.-J., Wu X.-F.;
RL "Cloning and Expression of Osteoprotegerin from Homo sapiens.";
RL Sheng Wu Hua Xueh Yu Sheng Wu Li Hsueh Pao 31:680-684(1999).
DR EMBL; AF134187; AAF20168.1; -;
DR HSSP; P25942; 1CDF.
DR INTERPRO; IPR001368; -;
DR PFAM; PF00020; TNFR_C6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS00500; TNFR_NGFR_2; 2.
DR NON_TER 1
SQ SEQUENCE 372 AA; 42758 MW; F02527A5CD01CCD3 CRC64;

Query Match 33.4%; Score 257; DB 4; Length 372;
Best Local Similarity 40.4%; Pred. No. 7.le-19;
Matches 40; Conservative 18; Mismatches 41; Indels 0; Gaps 0;

QY 34 PTYPWRDAETGRLVCAACPPGTFVQRCRDRSPPTTCGCPPRHYTQFWNYLERCYCNV 93
DB 5 PKYLHYDEETSHQLLCKDCPCPTLYLKQHCCTAKWKIVCAPCPDHYTDSWHTSDEGLYCSP 64

QY 94 LCGEEREARACHATHNACRCRTGFFAHAGFCLEHASC 132
DB 65 VCKELQYVQKQCNRTHNRVYCECKEGRIEIEFCLKHRSC 103

RESULT 4
O08712 PRELIMINARY; PRT; 401 AA.
AC O08712; O70202;
DT 01-JUL-1997 (TRENBLrel. 04, Created)
DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR) (OCIF).
GN TNFRSF11B OR OPG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID-10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=KIDNEY;
RA MEDLINE-97262071; PubMed-9108485;
RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S., Luethy R., Nguyen H.O., Woodson S., Bennett L., Boone T., Shimamoto G., Derose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J., Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W., Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R., Suggs S., Boyle W.J.;
RA "Osteoprotegerin: a novel secreted protein involved in the regulation of bone density.";
RT Cell 89:309-319(1997).
RL [2]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/OLA, AND NIH SWISS;
RX MEDLINE-98382527; PubMed-9714833;
RA Mizuno A., Murakami A., Nakagawa N., Yasuda H., Tsuda E., Morinaga T., Higashio K.;

RT "Structure of the mouse osteoclastogenesis inhibitory factor (OCIF) gene and its expression in embryogenesis.";
RL Gene 215:339-343(1998).
CC -!- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN STROMAL CELLS AND OSTEOCLAST PROGENITORS.
CC -!- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM).
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN LIVER, LUNG, BRAIN, HEART, KIDNEY, STOMACH, INTESTINE, SKIN, CALVARIA AND PLACENTA. NOT DETECTED IN SPLEEN.
CC -!- DEVELOPMENTAL STAGE: IN THE EMBRYO, HIGH LEVELS ARE DETECTED AT DAY 7. AT DAY 11, EXPRESSION DECREASES AND THEN INCREASES FROM DAY 15 TO DAY 17.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL; U94331; AAB53708.1; -;
DR EMBL; AB013898; BAA28269.1; -;
DR EMBL; AB013903; BAA33388.1; -;
DR EMBL; AB013899; BAA33388.1; JOINED.
DR EMBL; AB013900; BAA33388.1; JOINED.
DR EMBL; AB013901; BAA33388.1; JOINED.
DR EMBL; AB013902; BAA33388.1; JOINED.
DR HSSP; P25942; 1CDF.
DR MGD; MGI:109587; Opg.
DR INTERPRO; IPR000488; -;
DR INTERPRO; IPR001368; -;
DR PFAM; PF00020; TNFR_C6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS00500; DEATH_DOMAIN; 1.
DR PROSITE; PS00500; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -; 1
KW Glycoprotein; Repeat; Cytokine; Signal.
FT SIGNAL 1
FT CHAIN 22 401 OSTEOPROTEGERIN.
FT DOMAIN 23 201 4 X TNFR-CYS.
FT REPEAT 23 63 TNFR-CYS 1.
FT REPEAT 64 106 TNFR-CYS 2.
FT REPEAT 107 143 TNFR-CYS 3.
FT REPEAT 144 201 TNFR-CYS 4.
FT DOMAIN 306 365 DEATH DOMAIN.
FT DISULFID 41 54 BY SIMILARITY.
FT DISULFID 44 62 BY SIMILARITY.
FT DISULFID 65 80 BY SIMILARITY.
FT DISULFID 83 97 BY SIMILARITY.
FT DISULFID 87 105 BY SIMILARITY.
FT DISULFID 118 142 BY SIMILARITY.
FT DISULFID 145 160 BY SIMILARITY.
FT CARBOHYD 98 98 POTENTIAL.
FT CARBOHYD 165 165 POTENTIAL.
FT CARBOHYD 178 178 POTENTIAL.
FT CARBOHYD 289 289 POTENTIAL.
FT VARIANT 138 138 R -> P (IN STRAINS 129/OLA AND NIH SWISS).
FT VARIANT 161 161 I -> R (IN STRAINS 129/OLA AND NIH SWISS).
FT VARIANT 165 165 N -> D (IN STRAINS 129/OLA AND NIH SWISS).
FT VARIANT 288 288 S -> A (IN STRAINS 129/OLA AND NIH SWISS).
FT VARIANT 296 296 L -> R (IN STRAINS 129/OLA AND NIH SWISS).
SQ SEQUENCE 401 AA; 45923 MW; CAA6102D3B312470 CRC64;

Query Match 32.4%; Score 246; DB 11; Length 401;
Best Local Similarity 40.4%; Pred. No. 1e-17;
Matches 40; Conservative 16; Mismatches 43; Indels 0; Gaps 0;

QY 34 PTYPWRDAETGRLVCAACPPGTFVQRCRDRSPPTTCGCPPRHYTQFWNYLERCYCNV 93
DB 26 PKYLHYDPETGHLQLLCKDCPCPTLYLKQHCCTAKWKIVCAPCPDHYTDSWHTSDEGLYCSP 85

Qy	94	LCGEREEARACHATHNRACRCRTGFFAHAGFCLBHASC	132
		: : :	
Db	86	VCKELQSVKQECNRTNRRYCECEEGRYLETFCLKHRSC	124

RESULT	5
ID	008727
AC	PRELIMINARY; PRT: 401 AA.
DT	008727;
DT	01-JUL-1997 (T-EMBLrel. 04, Created)
DT	01-JUL-1997 (T-EMBLrel. 04, Last sequence update)
DT	01-OCT-2000 (T-EMBLrel. 15, Last annotation update)
DE	OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR) (OCIF).
GN	TNFRSF11B OR OPG.
OS	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rata
OX	NCBI_TaxID=101116;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=INTESTINE;
RC	MEDLINE=97262071; PubMed=9108485;
RA	Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,
RA	Luethy R., Nguyen H.O., Wooden S., Bennett L., Boone T., Shimamo
RA	Derosé M., Elliott R., Colombero A., Tan H.-L.; Trail G., Sullivan
RA	Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pettis
RA	Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,
RA	Suggs S., Boyle W.J.;
RT	"Osteoprotegerin: a novel secreted protein involved in the regul
RL	Cell 89:309-319(1997).
CC	-I- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF N SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETW STROMAL CELLS AND OSTEOCLAST PROGENITORS. -I- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM) (BY SIMILARITY).
CC	-I- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC	-I- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION
DR	EMBL: U94330; AAB53707.1; -
DR	HSSP: P25942; ICDF
DR	INTERPRO: IPR001368; -
DR	PFAM: PF000020; TNFR_c6; 4.
DR	PROSITE: PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR	PROSITE: PS50050; TNFR_NGFR_2; 2.
DR	PRODOM: PD000771; -; 1.
KW	Glycoprotein; Repeat; Cytokine; Signal.
FT	SIGNAL 1 21 BY SIMILARITY.
FT	CHAIN 22 401 OSTEOPROTEGERIN.
FT	DOMAIN 23 201 4 X TNFR-CYS.
FT	REPEAT 23 63 TNFR-CYS 1.
FT	REPEAT 64 106 TNFR-CYS 2.
FT	REPEAT 107 143 TNFR-CYS 3.
FT	REPEAT 144 201 TNFR-CYS 4.
FT	DOMAIN 306 365 DEATH DOMAIN.
FT	DISULFID 41 54 BY SIMILARITY.
FT	FT DISULFID 44 62 BY SIMILARITY.
FT	FT DISULFID 65 80 BY SIMILARITY.
FT	FT DISULFID 83 97 BY SIMILARITY.
FT	FT DISULFID 87 105 BY SIMILARITY.
FT	FT DISULFID 118 142 BY SIMILARITY.
FT	FT DISULFID 145 160 BY SIMILARITY.
FT	FT CARBOHYD 98 98 POTENTIAL.
FT	FT CARBOHYD 165 165 POTENTIAL.
FT	FT CARBOHYD 178 178 POTENTIAL.
FT	FT CARBOHYD 289 289 POTENTIAL.
SQ	SEQUENCE 401 AA; 46192 MW; FEC6A31FD4E573A CRC64;

Query Match 32.0%; Score 243; DB 11; Length 401;

	Best Local Similarity	40.48;	Pred. No.	2.le-17;	
	Matches	40;	Conservative	16;	Mismatches 43; Indels 0; Gaps 0;
QY	34	TPYPWRDAETGERLVCAOCPPGTFVORPCRRDSPPTCGCPPRHYTQFWNYLERCRYCNV	93		
		: : : : : : : : : : : : : : : :			
Db	26	PXYLHYDETGRQLLKCDKACGTLYKHCHCTVRRRTLVCVPDYPDYSYTSWHSDECVYCSP	85		
		: : : : : : : : : : : : : : : :			
QY	94	LGEEREERARACHATHNRACRCRTGFFFAHAGFCLEHASC	132		
		: : : : : : : : : : : : : : : : : : : :			
Db	86	VCKELQTVKQCNRTHNRVCECEEGRYLELEFCLKHRSC	124		
		: : : : : : : : : : : : : : : :			
RESULT	6				
ID	O9PUSO	PRELIMINARY:	PRT;	302 AA.	
AC	O9PUSO;				
DT	01-MAY-2000 (TREMBLrel. 13, Created)				
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)				
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)				
DE	DECAY RECEPTOR.				
OS	Salvelinus fontinalis (Brook trout).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei;				
OC	Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.				
OX	NCBI_Taxid=8038;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Bobe J., Goetz F.W.;				
RT	"A tumor necrosis factor receptor homolog is up-regulated in the brook				
RL	trout (Salvelinus fontinalis) ovary at the completion of ovulation.";				
DR	Biol. Reprod. 0:0-0(1999).				
DR	EMBL; AF156738; AAD56428.1; -.				
DR	HSSP; P19438; 1EXT.				
DR	INTERPRO; IPR000561; -.				
DR	INTERPRO; IPR001368; -.				
DR	PFAM; PF00020; TNFR_C6; 4.				
DR	PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.				
DR	PROSITE; PS01186; EGF_2; UNKNOWN_1.				
DR	PROSITE; PS50050; TNFR_NGFR_2; 1.				
KW	Receptor.				
SQ	SEQUENCE 302 AA; 34037 MW; E44C73477F05C3DF CRC64;				
	Query Match	26.9%;	Score	204.5;	DB 13; Length 302;
	Best Local Similarity	32.0%;	Pred. No.	1.5e-13;	
	Matches	32;	Conservative	25;	Mismatches 40; Indels 3; Gaps 2;
QY	35	TPYPWRDAETGERLVCAOCPPGTFVORPCRRDSPPTCGCPPRHYTQFWNYLERCRYCNV	94		
		: : : : : : : : : : : : : : : : : : :			
Db	22	TFKNDTRYSGLSIVDCRCPGTYLRAPCSAMRKSDCAECPNGAYTERWNHISKLRCS-M	80		
		: : : : : : : : : : : : : : : :			
QY	95	CGEEREERARACHATHNRACRCRTGFFFAHAGF-CLEHASC	132		
		: : : : : : : : : : : : : : : :			
Db	81	CAENOVVKEGSPSNCECKEGYFNKKYEACIKHKEC	120		
		: : : : : : : : : : : : : : : :			
RESULT	7				
ID	O57277	PRELIMINARY:	PRT;	348 AA.	
AC	O57277;				
DT	01-JUN-1998 (TREMBLrel. 06, Created)				
DT	01-JUN-1998 (TREMBLrel. 06, Last sequence update)				
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)				
DE	TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.				
GN	CRMB.				
OS	Monkeypox virus.				
OC	Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;				
OC	Orthopoxvirus.				
OX	NCBI_Taxid=10244;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-ZAIRE-1996(96-17), ZAIRES-1996 (96-16);				
RA	Loparev V.N., Parsons J.M., Esposito J.J.;				

```

RESULT 10
OS7108 PRELIMINARY; PRT; 348 AA.
ID OS7108
AC OS7108;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN CRMB.
OS Monkeypox virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae.
OC Orthopoxvirus.
OX NCBI_TaxID=10244;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZAIRE-1970 (CONGO-8);
RA Loparev V.N., Parsons J.M., Esposito J.J.;
RL Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL; U88142; AAB94367.1; -.
DR HSSP; P25942; 1CDF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 2.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -.
DR PRODOM; PD000771; 1.
SQ SEQUENCE 348 AA; 38212 MW; E555979057DEC91F CRC64;

```


Db 10 LFLSCIIINGRDIAHAPSNGKCKDNEYRSN-----LCCLSCPPGTYASRLCDSKTNT 63
Qy 69 TCGPCPPRHYTFQWNYLRCRYCNVLCGEREEARACHATHNRACRCRTGFF-----AH 122
Db 64 QCTPCGSDTFTSHNHQLQACLSGCRCDNSQVETRSCNTTHNRICECSPGYICLLKGASG 123
Qy 123 AGFCLHASC 132
Db 124 CRTCSKTKC 133
RESULT 14
O57099
ID O57099 PRELIMINARY; PRT; 349 AA.
AC O57099;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN CRMB.
OS Monkeypox virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10244;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SIERRA LEONE-1970 (70-0266);
RA Loparev V.N., Parsons J.M., Esposito J.J.;
RL Submitted (JAN-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL; U87843; AAB94360.1; -.
DR HSP; P25942; ICDF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 2.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -.
SQ SEQUENCE 349 AA; 38321 MW; FE449028CC933F57 CRC64;

Query Match 25.4%; Score 192.5; DB 12; Length 349;
Best Local Similarity 32.8%; Pred. No. 3e-12;
Matches 39; Conservative 16; Mismatches 53; Indels 11; Gaps 2;
Qy 9 LSLCLVLALPALLPVPAVRGVAETPTVPWRDAETGERLVAQCQPPGTFVORPCRRDSPT 68
Db 10 LFLSCIIINGRDIAHAPSNGKCKDNEYRSN-----LCCLSCPPGTYASRLCDSKTNT 63
Qy 69 TCGPCPPRHYTFQWNYLRCRYCNVLCGEREEARACHATHNRACRCRTGFFAHAGFCL 127
Db 64 QCTPCGSDTFTSHNHQLQACLSGCRCDNSQVETRSCNTTHNRICECSPGY-----YCL 117

RESULT 15
Q62327
ID Q62327 PRELIMINARY; PRT; 459 AA.
AC Q62327;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR 2 MRNA (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD;
RX MEDLINE=95178848; PubMed=7873884;
RA Powell E.E., Wicker L.S., Peterson L.B., Todd J.A.;
RT "Allelic variation of the type 2 tumor necrosis factor receptor
gene.";
RL Mamm. Genome 5:726-727(1994).
DR EMBL; X76401; CAA53981.1; -.

DR HSP; P19438; INCF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00050; TNFR_NGFR_2; 3.
FT NON_TER 1
FT VARIANT 87 87 S -> T.
FT VARIANT 93 93 T -> I.
FT VARIANT 268 268 F -> I.
FT VARIANT 345 345 S -> F.
FT VARIANT 421 421 Y -> C.
SQ SEQUENCE 459 AA; 48686 MW; 6C51D2CF1C4626DF CRC64;
Query Match 24.9%; Score 189; DB 11; Length 459;
Best Local Similarity 37.6%; Pred. No. 8.8e-12;
Matches 32; Conservative 10; Mismatches 39; Indels 4; Gaps 1;
Qy 46 RLVAQCQPPGTFVORPCRRDSPTTCGPPRHYTFQWNYLRCRYCNVLCGEREEARAC 105
Db 37 QMCCAKCPPGGYVVKHFCNKTSDTVCADCEASMYTQVWVQFRTCLSCSSCSTDQVETRAC 96
Qy 106 HATHNRACRCRTGFF-----AHAGFC 126
Db 97 TKQONRVCAACEAGRYCALKTHSGSC 121
Search completed: January 30, 2001, 16:54:58
Job time: 914 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:47:47 ; Search time 132.77 Seconds
(without alignments)
17.953 Million cell updates/sec

Title: US-09-518-931-4_COPY_1_132

Perfect score: 759

Sequence: 1 MRALEGPGLSLCLVLALPA.....CRCRTGFFAHAGFCLEHASC 132

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6.COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/PCRUS.COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	759	100.0	300	2	US-08-794-796-2
2	260.5	34.3	401	3	US-08-974-022-6
3	246	32.4	401	3	US-08-974-022-4
4	243	32.0	401	3	US-08-974-022-2
5	186	24.5	474	2	US-08-650-000-4
6	186	24.5	474	5	5395760-4
7	181	23.8	355	1	US-08-292-549-6
8	176	23.2	197	2	US-08-505-606-1
9	176	23.2	227	3	US-08-974-022-48
10	176	23.2	461	1	US-08-385-229-2
11	176	23.2	461	2	US-08-650-000-2
12	176	23.2	461	5	5395760-2
13	176	23.2	486	1	US-08-243-010-1
14	176	23.2	518	1	US-08-385-229-4
15	168	22.1	120	3	US-08-974-022-42
16	168	22.1	655	3	US-08-959-382-2
17	165	21.7	163	2	US-08-219-237B-5
18	163	21.5	164	2	US-08-232-087A-9
19	152	20.0	326	1	US-08-292-549-4
20	152	20.0	326	4	PCT-US91-02207-4
21	151.5	20.0	207	3	US-08-974-022-47
22	151.5	20.0	325	1	US-08-292-549-2
23	151.5	20.0	325	4	PCT-US91-02207-2
24	141.5	18.6	277	2	US-08-147-784-2
25	140	18.4	186	1	US-08-089-458B-6
26	133.5	17.6	283	4	PCT-US96-12374-2
27	124	16.3	125	3	US-08-959-382-4
28	123	16.2	70	3	US-08-974-022-41

Sequence 11, Appl
Sequence 51, Appl
Sequence 15, Appl
Sequence 8, Appl
Sequence 3, Appl
Sequence 7, Appl
Sequence 45, Appl
Sequence 19, Appl
Sequence 19, Appl
Sequence 4, Appl
Sequence 2, Appl
Sequence 1, Appl
Sequence 2, Appl
Sequence 11, Appl
Sequence 11, Appl
Sequence 43, Appl

ALIGNMENTS

RESULT 1
US-08-794-796-2
: Sequence 2, Application US/08794796
: Patent No. 5885800
: GENERAL INFORMATION:
: APPLICANT: Emery, John
: APPLICANT: Tan, KB
: APPLICANT: Truneh, Alem
: APPLICANT: Young, Peter
: TITLE OF INVENTION: Tumor Necrosis Related Receptor.
: TITLE OF INVENTION: TR4
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SmithKline Beecham Corporation
: STREET: 709 Swedeland Road
: CITY: King of Prussia
: STATE: PA
: COUNTRY: USA
: ZIP: 19406
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/794,796
: FILING DATE: 04-FEB-1997
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Han, William T
: REGISTRATION NUMBER: 34,344
: REFERENCE/DOCKET NUMBER: GH50000
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 610-270-5219
: TELEFAX: 610-270-4026
: TELEX:
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 300 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-794-796-2

Query Match 100.0%; Score 759; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 1.1e-64;


```
;
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,022
; FILING DATE: 12-DEC-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/577,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-378
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 401 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-974-022-2

Query Match 32.0%; Score 243; DB 3; Length 401;
Best Local Similarity 40.4%; Pred. No. 9.3e-16;
Matches 40; Conservative 16; Mismatches 43; Indels 0; Gaps 0;

QY 34 PTYPMRDAETGERLVCAQCPGTFVQPCRRDSTTCGPPRHYTQFWNYLERCRYCNV 93
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 26 PYLHYDPETGQLLCKDKCAPTYLKHCTVRRKTLCPDPSYSDSWHTSDECVYCSP 85
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 94 LCGEEREARACHATHNRACRRTGFFAHAGFCLEHASC 132
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 86 VCKELQTVKQECNRTHNRVCECEGGRYLEFCLKHRSC 124
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

RESULT 5
US-08-650-000-4
; Sequence 4, Application US/08650000
; Patent No. 5945397
; GENERAL INFORMATION:
; APPLICANT: Smith, Craig A.
; APPLICANT: Goodwin, Raymond G.
; APPLICANT: Beckmann, M. Patricia
; TITLE OF INVENTION: Tumor Necrosis Factor Receptors
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/650,000
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,453
; FILING DATE:
; APPLICATION NUMBER: US/08/038,765
; FILING DATE:
; APPLICATION NUMBER: US 403,241
; FILING DATE: 05-SEP-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 405,370
```

```
;
; FILING DATE: 11-SEP-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 421,417
; FILING DATE: 13-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 523,635
; FILING DATE: 10-MAY-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Wight, Christopher L.
; REGISTRATION NUMBER: 31,680
; REFERENCE/DOCKET NUMBER: 2501-D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 474 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-650-000-4

Query Match 24.5%; Score 186; DB 2; Length 474;
Best Local Similarity 37.6%; Pred. No. 2.7e-10;
Matches 32; Conservative 10; Mismatches 39; Indels 4; Gaps 1;

QY 46 RLVCAQCPGTFVQPCRRDSTTCGPPRHYTQFWNYLERCRYCNVLCGEREEARAC 105
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 52 QMCKAKCPGQGVKHFCKNTSDTVACDEASMYTQVWNOFRTCLSCSSCTTDDQVEIRAC 111
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 106 HATHNRACRRTGFF---AHAGFC 126
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 112 TKQNRVCAEAGRYCALKTHSGSC 136
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

RESULT 6
5395760-4
; Patent No. 5395760
; APPLICANT: SMITH, CRAIG A.;GOODWIN, RAYMOND G.;BECKMANN,
; M. PATRICIA
; TITLE OF INVENTION: DNA ENCODING TUMOR NECROSIS FACTOR-a AND
; B-RECEPTORS
; NUMBER OF SEQUENCES: 17
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/523,635
; FILING DATE: 10-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 421,417
; FILING DATE: 13-OCT-1989
; APPLICATION NUMBER: 405,370
; FILING DATE: 11-SEP-1989
; APPLICATION NUMBER: 403,241
; FILING DATE: 05-SEP-1989
; SEQ ID NO:4:
; LENGTH: 474
; 5395760-4

Query Match 24.5%; Score 186; DB 5; Length 474;
Best Local Similarity 37.6%; Pred. No. 2.7e-10;
Matches 32; Conservative 10; Mismatches 39; Indels 4; Gaps 1;

QY 46 RLVCAQCPGTFVQPCRRDSTTCGPPRHYTQFWNYLERCRYCNVLCGEREEARAC 105
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 52 QMCKAKCPGQGVKHFCKNTSDTVACDEASMYTQVWNOFRTCLSCSSCTTDDQVEIRAC 111
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 106 HATHNRACRRTGFF---AHAGFC 126
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 112 TKQNRVCAEAGRYCALKTHSGSC 136
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

RESULT 7
```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: James F. Haley, Jr.
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10020

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/505,606
FILING DATE: 21-JUL-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/378,968
FILING DATE: 26-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: HALEY, Jr., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: B191
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 596-9000
TELEFAX: (212) 596-9090
TELEX: 14-8367
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 197 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-505-606-1

Query Match 23.28; Score 176; DB 2; Length 197;
Best Local Similarity 37.28; Pred. No. 9.5e-10; Gaps
Matches 35; Conservative 11; Mismatches 42; Indels 6; Gaps

QY 42 ETGERLVQAQPGFGVQRCRRDSPTCTGCPRRHYTQFWNYL---ERCRCYNVLGGER 98
Db 25 EPOHRCRCRCPGGYVSAKCSRIKDTVCATCAENSHNEHWNYLTICQLCRPCDPVWG-- 82

QY 99 EEPARACHATHNRACRRTGTFGAHAGCFLHASC 132
Db 83 LEEIAPTCKRRKTQCRCPGMEC-AAWALECTHC 115

RESULT 9
US-08-974-022-48
Sequence 48, Application US/08974022
Patent No. 6015938
GENERAL INFORMATION:
APPLICANT: Boyle, William J.
APPLICANT: Lacey, David L.
APPLICANT: Calzone, Frank J.
APPLICANT: Chang, Ming-Shi
TITLE OF INVENTION: OSTEOPROTEGERIN
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:45:29 ; Search time 183.67 seconds
(without alignments)
4.468 Million cell updates/sec

Title: US-09-518-931-4_COPY_57_80

Perfect score: 148

Sequence: 1 FVQPCRRDSTTCGCPPRHYTQ 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_36:*

1: /SIDSL1/gcgdata/geneseq/geneseq/AA1980.DAT:*

2: /SIDSL1/gcgdata/geneseq/geneseq/AA1981.DAT:*

3: /SIDSL1/gcgdata/geneseq/geneseq/AA1982.DAT:*

4: /SIDSL1/gcgdata/geneseq/geneseq/AA1983.DAT:*

5: /SIDSL1/gcgdata/geneseq/geneseq/AA1984.DAT:*

6: /SIDSL1/gcgdata/geneseq/geneseq/AA1985.DAT:*

7: /SIDSL1/gcgdata/geneseq/geneseq/AA1986.DAT:*

8: /SIDSL1/gcgdata/geneseq/geneseq/AA1987.DAT:*

9: /SIDSL1/gcgdata/geneseq/geneseq/AA1988.DAT:*

10: /SIDSL1/gcgdata/geneseq/geneseq/AA1989.DAT:*

11: /SIDSL1/gcgdata/geneseq/geneseq/AA1990.DAT:*

12: /SIDSL1/gcgdata/geneseq/geneseq/AA1991.DAT:*

13: /SIDSL1/gcgdata/geneseq/geneseq/AA1992.DAT:*

14: /SIDSL1/gcgdata/geneseq/geneseq/AA1993.DAT:*

15: /SIDSL1/gcgdata/geneseq/geneseq/AA1994.DAT:*

16: /SIDSL1/gcgdata/geneseq/geneseq/AA1995.DAT:*

17: /SIDSL1/gcgdata/geneseq/geneseq/AA1996.DAT:*

18: /SIDSL1/gcgdata/geneseq/geneseq/AA1997.DAT:*

19: /SIDSL1/gcgdata/geneseq/geneseq/AA1998.DAT:*

20: /SIDSL1/gcgdata/geneseq/geneseq/AA1999.DAT:*

21: /SIDSL1/gcgdata/geneseq/geneseq/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	148	100.0	170	W63623	Human tumour necro
2	148	100.0	245	Y28449	A human tumour nec
3	148	100.0	271	Y42184	Human mFLINT #1 pr
4	148	100.0	273	Y42185	Human mFLINT #2 pr
5	148	100.0	300	W66102	Amino acid sequenc
6	148	100.0	300	W63622	Human tumour necro
7	148	100.0	300	Y03099	Human lung TNF-rec
8	148	100.0	300	Y42182	Human FLINT #1 pro
9	148	100.0	300	Y17479	Mammalian tumour n
10	148	100.0	300	Y06817	Human Dcr3 polypep
11	148	100.0	300	W97749	Human tumour necro
12	148	100.0	300	W95082	Orphan receptor (H

13	148	100.0	302	20	Y42183	Human FLINT #2 pro
14	142	95.9	300	21	Y77458	Human TNF receptor
15	56	37.8	84	17	R99951	Mutated OCIF, OCIF
16	56	37.8	106	17	R99947	Mutated OCIF, OCIF
17	56	37.8	143	17	R99946	Mutated OCIF, OCIF
18	56	37.8	145	17	R99930	Osteoclastogenesis
19	56	37.8	154	17	R99929	Osteoclastogenesis
20	56	37.8	187	17	R99950	Mutated OCIF, OCIF
21	56	37.8	187	21	Y77464	Primate protein se
22	56	37.8	197	17	R99945	Mutated OCIF, OCIF
23	56	37.8	272	17	R99944	Mutated OCIF, OCIF
24	56	37.8	321	17	R99949	Mutated OCIF, OCIF
25	56	37.8	326	17	R99940	Mutated OCIF, OCIF
26	56	37.8	327	17	R99941	Mutated OCIF, OCIF
27	56	37.8	349	20	W83928	Human FTHMA-070 pa
28	56	37.8	351	17	R99943	Mutated OCIF, OCIF
29	56	37.8	359	17	R99939	Mutated OCIF, OCIF
30	56	37.8	360	17	R99938	Mutated OCIF, OCIF
31	56	37.8	380	17	R99924	Mature osteoclasto
32	56	37.8	390	17	R99357	Human tumour necro
33	56	37.8	391	19	W53238	Human OCIF genome
34	56	37.8	393	17	R99948	Mutated OCIF, OCIF
35	56	37.8	395	19	W57636	Modified TRI recep
36	56	37.8	399	17	R99942	Mutated OCIF, OCIF
37	56	37.8	401	17	R99932	Mutated OCIF, OCIF
38	56	37.8	401	17	R99933	Mutated OCIF, OCIF
39	56	37.8	401	17	R99934	Mutated OCIF, OCIF
40	56	37.8	401	17	R99935	Mutated OCIF, OCIF
41	56	37.8	401	17	R99925	Full length osteoc
42	56	37.8	401	17	R99931	Mutated OCIF, OCIF
43	56	37.8	401	18	W38345	Human osteoprotege
44	56	37.8	401	19	W57635	TRI receptor prote
45	56	37.8	401	20	W83926	Human FTHMA-070 pr

ALIGNMENTS

RESULT	1
W63623	ID W63623 standard; Protein; 170 AA.
XX	XX
AC	W63623;
XX	XX
DT	26-OCT-1998 (first entry)
XX	XX
DE	Human tumour necrosis factor receptor-6 beta protein.
XX	XX
KW	Human tumour necrosis factor receptor-6 alpha; TNFR-6 alpha; TNFR-6 beta;
KW	endothelial cells; keratinocytes; normal prostate; apoptosis;
KW	prostate tumour tissue.
XX	XX
OS	Homo sapiens.
XX	XX
FH	Key
FT	Peptide
FT	Protein
FT	Location/Qualifiers
FT	1..30
FT	31..170
FT	note="TNFR-6 beta"
FT	31..166
FT	/note="Soluble extracellular domain"
XX	XX
PN	W09830694-A2.
XX	XX
PD	16-JUL-1998.
XX	XX
PF	13-JAN-1998; 98WO-US00153.
XX	XX
PR	14-JAN-1997; 97US-0035496.
XX	XX
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	XX
PI	Ebner R, Feng P, Gentz RL, Ni J, Ruben SM, Yu G;
XX	XX

```

DR WPI; 1998-399142/34.
DR N-PSDB; V39086.
XX
PT Human tumour necrosis factor receptors 6-alpha and 6-beta - used in
PT the diagnosis of immune system-related disorder(s)
XX
XX Claim 20; Fig 2A-2B; 9lpp; English.
XX
CC The present sequence represents the human tumour necrosis factor
CC receptor-6 beta (TNFR-6 beta) protein. The invention also provides
CC for the TNFR-6 alpha protein (W63622). TNFR-6 alpha and TNFR-6 beta are
CC members of the tumour necrosis factor receptor (TNFR) family. TNFRs
CC are expressed in endothelial cells, keratinocytes, normal prostate and
CC prostate tumour tissue. For a number of disorders of these cells,
CC particularly of the immune system, substantially altered (whether
CC increased or decreased) levels of TNFR-6 alpha and/or TNFR-6 beta gene
CC expression can be detected, therefore the TNFR-6 alpha and TNFR-6 beta
CC polypeptides, nucleic acids and antibodies are claimed to be useful in
CC the diagnosis of such disorders. Mutations of the TNFR-6 alpha and
CC TNFR-6 beta genes can also be detected. The TNFR polypeptides are
CC also claimed to be useful for identifying ligands which may be useful
CC in the treatment of apoptosis related disorders.
XX
SQ Sequence 170 AA;
Query Match 100.0%; Score 148; DB 19; Length 170;
Best Local Similarity 100.0%; Pred. No. 2.7e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVQPCRRDSPTTCGCPPRHYTQ 24
DB 57 fvqpcrrdspttcgpcprhytq 80

RESULT 2
Y28449
ID Y28449 standard; Protein; 245 AA.
XX
AC Y28449;
XX
DT 29-SEP-1999 (first entry)
XX
DE A human tumour necrosis factor-R2-like proteins (TR2P)-1.
XX
KW Human tumour necrosis factor-R2-like protein; TR2P; achondroplasia;
KW osteoporosis; developmental disorder; Cushing's syndrome;
KW muscular dystrophy; epilepsy; hereditary neuropathy;
KW Charcot-Marie-Tooth disease; neurofibromatosis; hypothyroidism;
KW hydrocephalus; seizure disorder; cerebellar palsy; spinal bifida;
KW congenital glaucoma; cataract; sensorineural hearing loss;
KW reproductive disorder; infertility; ovulatory defect; endometriosis;
KW autoimmune disorder; ectopic pregnancy; teratogenesis; spermatogenesis;
KW immunological disorder; AIDS; Addison's disease; allergy; bronchitis;
KW atherosclerosis; diabetes mellitus; Chron's disease; lupus;
KW irritable bowel syndrome; multiple sclerosis; infection;
KW neoplastic disorder; adenocarcinoma; leukaemia; lymphoma;
KW myeloma; sarcoma.
XX
OS Homo sapiens.
XX
PN WO9931128-A2.
XX
PD 24-JUN-1999.
XX
PF 02-DEC-1998; 98WO-US25649.
XX
PR 16-DEC-1997; 97US-0991945.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Au-Young J, Bandman O, Hillman JL, Kaser MR, Tang YT;
XX
WPI; 1999-457916/38.
DR N-PSDB; X89503.
XX
PT New tumour necrosis factor-R2-like protein - useful in the treatment
PT of osteogenesis, developmental, reproductive, immunological and
PT neoplastic disorders
XX
XX Claim 1; Fig 1A-C; 8lpp; English.
XX
CC The present sequence represents a human tumour necrosis factor-R2-like
CC protein (TR2P). The protein is used to treat and prevent osteogenesis,
CC developmental, reproductive, immunological and neoplastic disorders, and
CC also to diagnose disorders associated with TR2 protein expression. Such
CC disorders include osteogenesis disorders such as achondroplasia and
CC osteoporosis, developmental disorders such as Cushing's syndrome,
CC muscular dystrophy, and epilepsy, hereditary neuropathies such as
CC Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism,
CC hydrocephalus, seizure disorders such as cerebral palsy and spinal
CC bifida, congenital glaucoma, cataract, or sensorineural hearing loss,
CC reproductive disorders such as infertility, ovulatory defects and
CC endometriosis, autoimmune disorders, ectopic pregnancy and teratogenesis,
CC disruption of spermatogenesis, immunological disorders such as AIDS,
CC Addison's disease, allergies, bronchitis, atherosclerosis, diabetes
CC mellitus, Chron's disease, lupus and irritable bowel syndrome, multiple
CC sclerosis, viral, fungal, helminthic, parasitic and protozoal infections,
CC and neoplastic disorders including adenocarcinoma, leukaemia, lymphoma,
CC melanoma, myeloma, sarcoma, and teratocarcinoma.
XX
SQ Sequence - 245 AA;
Query Match 100.0%; Score 148; DB 20; Length 245;
Best Local Similarity 100.0%; Pred. No. 3.7e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVQPCRRDSPTTCGCPPRHYTQ 24
DB 57 fvqpcrrdspttcgpcprhytq 80

RESULT 3
Y42184
ID Y42184 standard; Protein; 271 AA.
XX
AC Y42184;
XX
DT 17-DEC-1999 (first entry)
XX
DE Human mFLINT #1 protein sequence.
XX
KW Human; FLINT; mFLINT; OPC3; tumour necrosis factor receptor; FasL;
KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
KW reperfusion-associated injury; aplastic anaemia; differentiation;
KW growth; myelodysplastic syndrome; pancytopenic condition;
KW myocardial ischaemia.
XX
OS Homo sapiens.
XX
PN WO9950413-A2.
XX
PD 07-OCT-1999.
XX
PF 30-MAR-1999; 99WO-US06797.
XX
PR 30-MAR-1998; 98US-0079856.
XX
PR 20-MAY-1998; 98US-0086074.
XX
PR 09-SEP-1998; 98US-0099643.
XX
PR 17-DEC-1998; 98US-0112577.
XX
PR 18-DEC-1998; 98US-0112703.
XX
PR 18-DEC-1998; 98US-0112933.
XX
PR 22-DEC-1998; 98US-0113407.
XX

```


PA (ELIL) LILLY & CO ELI.
 XX Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
 PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
 PI Song HY, Wang J, Wu X, Zuckerman SH;
 XX WPI; 1999-591319/50.
 DR N-PSDB; 225377.
 XX Use of mature FLINT for treating acute liver failure, inflammation,
 PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
 PT and proinflammatory activity
 XX Claim 31; Fig 3; 99pp; English.
 XX The present invention describes therapeutic applications of mature FLINT
 CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
 CC (mFLINT), which is a member of the tumor necrosis factor receptor
 CC superfamily, is used for treating acute liver failure, inflammation of
 CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
 CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
 CC injury or disorder such as hypercoagulation (including use with
 CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
 CC or disorder, Type I diabetes, cancer, cell damage or damage to an
 CC innocent bystander tissue that is induced by a chemotherapeutic agent or
 CC therapeutic irradiation, treating haematopoietic progenitor cells that
 CC have been exposed to therapeutic radiation or chemotherapy, aplastic
 CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
 CC also used for promoting the growth or differentiation of a haematopoietic
 CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
 CC resulting from abnormal myocardial ischaemia. The present sequence
 CC represents human mFLINT.
 XX Sequence 271 AA;
 SQ

Query Match 100.0%; Score 148; DB 20; Length 271;
 Best Local Similarity 100.0%; Pred. No. 4e-10;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVQPCRDSPPTCGPCPPRHQTQ 24
 DB 28 fvqpcrrdspttcgpcpprhytq 51

RESULT 4
 Y42185
 ID Y42185 standard; Protein; 273 AA.
 XX AC Y42185;

DT 17-DEC-1999 (first entry)

DE Human mFLINT #2 protein sequence.

KW Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
 KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
 KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
 KW reperfusion-associated injury; aplastic anaemia; differentiation;
 KW growth; myelodysplastic syndrome; pancytopenic condition;
 KW myocardial ischaemia.

XX Homo sapiens.

XX WO9950413-A2.

PN 07-OCT-1999.

XX 30-MAR-1999; 99WO-US06797.

XX 30-MAR-1998; 98US-0079856.

PR 20-MAY-1998; 98US-0086074.

PR 09-SEP-1998; 98US-0099643.

PR 17-DEC-1998; 98US-0112577.
 PR 18-DEC-1998; 98US-0112703.
 PR 18-DEC-1998; 98US-0112933.
 PR 22-DEC-1998; 98US-0113407.
 XX (ELIL) LILLY & CO ELI.
 XX Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
 PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
 PI Song HY, Wang J, Wu X, Zuckerman SH;
 XX WPI; 1999-591319/50.
 DR N-PSDB; 225378.
 XX Use of mature FLINT for treating acute liver failure, inflammation,
 PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
 PT and proinflammatory activity
 XX Example 2; Fig 4; 99pp; English.
 XX The present invention describes therapeutic applications of mature FLINT
 CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
 CC (mFLINT), which is a member of the tumor necrosis factor receptor
 CC superfamily, is used for treating acute liver failure, inflammation of
 CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
 CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
 CC injury or disorder such as hypercoagulation (including use with
 CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
 CC or disorder, Type I diabetes, cancer, cell damage or damage to an
 CC innocent bystander tissue that is induced by a chemotherapeutic agent or
 CC therapeutic irradiation, treating haematopoietic progenitor cells that
 CC have been exposed to therapeutic radiation or chemotherapy, aplastic
 CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
 CC also used for promoting the growth or differentiation of a haematopoietic
 CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
 CC resulting from abnormal myocardial ischaemia. The present sequence
 CC represents human mFLINT.
 XX Sequence 273 AA;
 SQ

Query Match 100.0%; Score 148; DB 20; Length 273;
 Best Local Similarity 100.0%; Pred. No. 4e-10;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVQPCRDSPPTCGPCPPRHQTQ 24
 DB 28 fvqpcrrdspttcgpcpprhytq 51

RESULT 5
 W66102
 ID W66102 standard; Protein; 300 AA.
 XX AC W66102;

DT 02-DEC-1998 (first entry)

DE Amino acid sequence of tumour necrosis related receptor (TR4).

KW Human; tumour necrosis related receptor; TR4; agonist; antagonist;
 KW inhibition; chronic; acute; inflammation; arthritis; septicemia;
 KW autoimmune disease; transplant rejection; stroke; cancer;
 KW Alzheimer's disease.

XX Homo sapiens.

XX EP861850-A1.

PN 02-SEP-1998.

XX 20-JAN-1998; 98EP-0300382.

```

PR 04-FEB-1997; 97US-0794796.
XX
XX (SMIK ) SMITHKLINE BEECHAM CORP.
XX
XX Emery J, Tan KB, Truneh A, Young PR;
XX
XX WPI; 1998-508248/44.
DR N-PSDB; V07654.
XX
XX New DNA encoding tumour necrosis related receptor - used to treat
XX and prevent e.g. inflammation, arthritis, septicemia, autoimmune
XX diseases, transplant rejection, infection, stroke, ischaemia, ARDS,
XX restenosis, AIDS, bone disorders and cancer
XX
XX Claim 1; Fig 1; 2ipp; English.
XX
XX This is the amino acid sequence of the human tumour necrosis related
XX receptor (TR4), used in the method of the invention. The TR4 protein
XX or its agonist can be used to treat a subject in need of enhanced
XX TR4 polypeptide activity. The antagonist is used to inhibit TR4
XX polypeptide activity. The active agents can be used for the
XX treatment and prevention of diseases such as chronic and acute
XX inflammation, arthritis, septicemia, autoimmune diseases, transplant
XX rejection, stroke, cancer, Alzheimer's disease.
XX
SQ Sequence 300 AA;

Query Match 100.0%; Score 148; DB 19; Length 300;
Best Local Similarity 100.0%; Pred. No. 4.4e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVQPCRRDSPTTCGCPPRHYTQ 24
Db 57 fvgpcrrdspttcgpcpprhytq 80
|||||
|||||

RESULT 6
W63622
ID W63622 standard; Protein; 300 AA.
XX
XX W63622;
XX
XX 26-OCT-1998 (first entry)
XX
XX Human tumour necrosis factor receptor-6 alpha protein.
XX
XX Human tumour necrosis factor receptor-6 alpha; TNFR-6 beta;
XX endothelial cells; keratinocytes; normal prostate; apoptosis;
XX prostate tumour tissue.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..30
XX Protein 31..300
XX Region /note= "TNFR-6 alpha"
XX /note= "Soluble extracellular domain"
XX
XX W09830694-A2.
XX
XX 16-JUL-1998.
XX
XX 13-JAN-1998; 98WO-US00153.
XX
XX 14-JAN-1997; 97US-0035496.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ebner R, Feng P, Gentz RL, Ni J, Ruben SM, Yu G;
XX WPI; 1998-399142/34.
XX
DR

DR N-PSDB; V39085.
XX
XX Human tumour necrosis factor receptors 6-alpha and 6-beta - used in
XX the diagnosis of immune system-related disorder(s)
XX
XX Claim 20; Fig 1; 9lpp; English.
XX
XX The present sequence represents the human tumour necrosis factor
XX receptor-6 alpha (TNFR-6 alpha) protein. The invention also provides
XX for the TNFR-6 beta protein (W63623). TNFR-6 alpha and TNFR-6 beta are
XX members of the tumour necrosis factor receptor (TNFR) family. TNFRs
XX are expressed in endothelial cells, keratinocytes, normal prostate and
XX prostate tumour tissue. For a number of disorders of these cells,
XX particularly of the immune system, substantially altered (whether
XX increased or decreased) levels of TNFR-6 alpha and/or TNFR-6 beta
XX expression can be detected, therefore the TNFR-6 alpha and TNFR-6 beta
XX polypeptides, nucleic acids and antibodies are claimed to be useful in
XX the diagnosis of such disorders. Mutations of the TNFR-6 alpha and
XX TNFR-6 beta genes can also be detected. The TNFR polypeptides are
XX also claimed to be useful for identifying ligands which may be useful
XX in the treatment of apoptosis related disorders.
XX
XX Sequence 300 AA;

Query Match 100.0%; Score 148; DB 19; Length 300;
Best Local Similarity 100.0%; Pred. No. 4.4e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVQPCRRDSPTTCGCPPRHYTQ 24
Db 57 fvgpcrrdspttcgpcpprhytq 80
|||||
|||||

RESULT 7
Y03099
ID Y03099 standard; Protein; 300 AA.
XX
XX Y03099;
XX
XX 09-DEC-1999 (first entry)
XX
XX Human lung TNF-receptor protein.
XX
XX Tumour necrosis factor; TNF; TNF receptor; human; lung; gene therapy;
XX detection; immunoassay; diagnosis; disease; immune system; tumour;
XX osteogenic system; cardiovascular system; central nervous system; asthma;
XX peripheral nervous systems; transplant incompatibility; antitumor;
XX rheumatoid arthritis; antiasthmatic; antiarthritic.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 134..1036
XX /*tag= a
XX /product= "TNF-receptor"
XX
XX DE19809978-A1.
XX
XX 16-SEP-1999.
XX
XX 09-MAR-1998; 98DE-1009978.
XX
XX 09-MAR-1998; 98DE-1009978.
XX
XX (BADI ) BASF AG.
XX
XX Kroeger B;
XX
XX WPI; 1999-519473/44.
XX
XX N-PSDB; Z09998.
XX
XX New soluble member of tumor necrosis factor receptor family, useful for

```

PT identification specific modulators and for treating disease e.g. tumors

XX
PS Claim 1; Page 8-9; 10pp; German.

XX This invention describes a novel tumour necrosis factor (TNF) receptor
CC (I) isolated from human lung tissue. (I) is used: (i) to raise specific
CC antibodies (Ab); (ii) to screen for specific (ant)agonists or ligands
CC (A), potential therapeutic agents; and (iii) therapeutically (optionally
CC expressed from a gene therapy vector) in conditions associated with a
CC deficit of (I). Ab are used: (a) for qualitative or quantitative
CC detection of (I) in standard immunoassays (for diagnosis of disease, or
CC susceptibility, or for monitoring); and (b) as therapeutic inhibitors in
CC cases where (I) is overexpressed. Nucleic acid (II) that encodes (I) is
CC used: (A) for recombinant production of (I); (B) also its oligonucleotide
CC fragments, in standard hybridization and/or amplification assays; (C) as
CC source of antisense molecules or ribozymes; and (D) to produce transgenic
CC animals (for studying (patho)physiology of (I)). Diseases possibly
CC associated with under- or over-expression of (I) are those of the immune,
CC osteogenic, cardiovascular and central or peripheral nervous systems,
CC tumors, transplant incompatibility, asthma and rheumatoid arthritis. The
CC products of the invention have antitumor, antiasthmatic and
CC antiarthritic activity. This sequence represents the TNF-receptor of the
CC invention.

XX Sequence 300 AA;

Query Match 100.0%; Score 148; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 4.4e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FVQPCRRDSTPTCGPPRHVTQ 24
Db 57 fvqpcrrdspttcgpcprhytq 80

RESULT 8
Y42182

ID Y42182 standard; Protein; 300 AA.

XX AC Y42182;

XX DT 17-DEC-1999 (first entry)

XX DE Human FLINT #1 protein sequence.

XX Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
KW reperfusion-associated injury; aplastic anaemia; differentiation;
KW growth; myelodysplastic syndrome; pancytopenic condition;
KW myocardial ischaemia.

XX OS Homo sapiens.

XX PN WO9950413-A2.

XX PD 07-OCT-1999.

XX PF 30-MAR-1999; 99WO-US06797.

XX PR 30-MAR-1998; 98US-0079856.

XX PR 20-MAY-1998; 98US-0086074.

XX PR 09-SEP-1998; 98US-0099643.

XX PR 17-DEC-1998; 98US-0112577.

XX PR 18-DEC-1998; 98US-0112703.

XX PR 22-DEC-1998; 98US-0112933.

XX PA (ELIL) LILLY & CO ELI.

XX PI Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;

PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
XX Song HY, Wang J, Wu X, Zuckerman SH;

XX DR WPI: 1999-591319/50.

XX DR N-PSDB; 225375.

XX Use of mature FLINT for treating acute liver failure, inflammation,
PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
PT and proinflammatory activity

XX Claim 30; Fig 1; 99pp; English.

XX The present invention describes therapeutic applications of mature FLINT
CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
CC (mFLINT), which is a member of the tumour necrosis factor receptor
CC superfamily, is used for treating acute liver failure. Inflammation of
CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
CC injury or disorder such as hypercoagulation (including use with
CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
CC or disorder, Type I diabetes, cancer, cell damage or damage to an
CC innocent bystander tissue that is induced by a chemotherapeutic agent or
CC therapeutic irradiation, treating haematopoietic progenitor cells that
CC have been exposed to therapeutic radiation or chemotherapy, aplastic
CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
CC also used for promoting the growth or differentiation of a haematopoietic
CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
CC resulting from abnormal myocardial ischaemia. The present sequence
CC represents human FLINT.

XX SQ Sequence 300 AA;

Query Match 100.0%; Score 148; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 4.4e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FVQPCRRDSTPTCGPPRHVTQ 24
Db 57 fvqpcrrdspttcgpcprhytq 80

RESULT 9

Y17479

ID Y17479 standard; Protein; 300 AA.

XX AC Y17479;

XX DT 02-AUG-1999 (first entry)

XX DE Mammalian tumour necrosis factor receptor OPG-2.

XX Tumour necrosis factor receptor; TNF receptor; OPG-2; Paget's disease;
KW osteopenic disorder; osteoclast activity; primary osteoporosis;
KW hyperglycaemia; osteolytic metastasis; immune response; cancer.

XX OS Mammalia.

XX PN WO9926977-A1.

XX PD 03-JUN-1999.

XX PF 24-NOV-1998; 98WO-US25065.

XX PR 17-FEB-1998; 98US-0074896.

XX PR 24-NOV-1997; 97US-0066446.

XX PA (BIOJ) BIOGEN INC.

XX PI Tschopp J;

XX DR WPI: 1999-347693/29.

XX DR N-PSDB; X76052.

XX New tumour necrosis factor family receptor OPG-2
 PT
 XX
 PS Claim 1: Page 18; 22pp; English.

XX The present sequence represents a mammalian tumour necrosis factor
 CC receptor, designated OPG-2. OPG-2, is a member of the tumour necrosis
 CC factor receptor family, and can be used: (i) to raise specific
 CC antibodies (Ab), (ii) to treat osteopenic disorders associated with
 CC excessive osteoclast activity, e.g. primary osteoporosis, Paget's
 CC disease, hyperglycaemia of malignancy, or osteolytic metastases; (iii)
 CC for affinity purification of cognate ligands, and (iv) to screen for
 CC ligands (antagonists or agonists). Ab, or other OPG-2 blocking agents
 CC such as soluble forms of the protein, are used to prevent, or reduce
 CC severity of, an immune response, and for treating cancer. They can also
 CC be used in diagnostic assays. The nucleic acid sequence encoding OPG-2
 CC can be used as a probe to isolate related sequences from other species.

XX Sequence 300 AA;

Query Match 100.0%; Score 148; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 4.4e-10;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVORPCRRDSTTCGPPRRHYTQ 24
 Db |||||
 57 fvgprcrrdspttcgpcpprrhytq 80

RESULT 10

Y06817
 ID Y06817 standard; Protein; 300 AA.

XX AC Y06817;

XX 24-JUN-1999 (first entry)

XX Human Dcr3 polypeptide.

XX Dcr3 polypeptide; tumour necrosis factor receptor; TNFR; Fas ligand;
 KW apoptosis; T cell mediated immune response; allergy; asthma; cancer;
 KW rheumatoid arthritis; Crohn's disease; guest vs. host disease; human;
 KW gene therapy.

XX Homo sapiens.

XX WO9914330-A1.

XX 25-MAR-1999.

XX 18-SEP-1998; 98WO-US19661.

XX 30-JUL-1998; 98US-0094640.

XX 18-SEP-1997; 97US-0059288.

XX (GETH) GENENTECH INC.

XX Ashkenazi AJ, Botstein D, Dodge KH, Goddard A, Gurney AL;
 PI Kim KU, Lawrence DA, Pitti R, Roy MA, Tumas DB;
 PI Wood WI;

XX WPI: 1999-244032/20.

XX N-PSDB; X32744.

XX Dcr3 polypeptide related to tumor necrosis factor receptor

XX Claim 5; Fig 1; 88pp; English.

XX This represents a human Dcr3 polypeptide, a homologue of tumour necrosis
 CC factor receptor (TNFR) polypeptide. Host cells containing a vector
 CC comprising the Dcr3 nucleic acid can be used for the recombinant
 CC expression of the protein. Dcr3 binds to Fas ligand, so it (or its

CC chimeras) are useful for modulating apoptosis in mammalian cells, also
 CC other Fas-ligand induced activities, particularly to inhibit T cell
 CC mediated immune responses, e.g. in treatment of allergy, asthma,
 CC rheumatoid arthritis, Crohn's disease, guest vs. host disease etc. Dcr3
 CC may also be used to identify specific binding proteins, potential
 CC inhibitors. Antibodies against Dcr3 are used to treat cancer,
 CC specifically of the lung and colon, also in diagnosis and for affinity
 CC purification of the protein. Detecting mutations in the gene for Dcr3 is
 CC also useful to diagnose cancer, or predisposition to it. Dcr3 nucleic acid
 CC is useful as hybridization probe to detect genomic or related sequences;
 CC for chromosome and gene mapping; as source of antisense sequences; for
 CC expression of recombinant Dcr3 and to generate transgenic animals (for
 CC development and screening of therapeutic agents), also for in vivo or
 CC ex vivo gene therapy.

XX Sequence 300 AA;

Query Match 100.0%; Score 148; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 4.4e-10;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVORPCRRDSTTCGPPRRHYTQ 24
 Db |||||
 57 fvgprcrrdspttcgpcpprrhytq 80

RESULT 11

W97749
 ID W97749 standard; Protein; 300 AA.

XX AC W97749;

XX 21-MAY-1999 (first entry)

XX Human tumour necrosis factor receptor ZTNFR-5.

XX ZTNFR-5; tumour necrosis factor receptor; TNFR; human;
 KW cell maturation; bone cell regulation.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..23 /note= "signal peptide"

FT Protein 24..300 /note= "mature protein"

FT Domain 24..194 /note= "extracellular domain"

FT Region 49..71 /note= "cysteine-rich pseudo-repeat 1"

FT Region 72..113 /note= "cysteine-rich pseudo-repeat 1"

FT Region 114..151 /note= "cysteine-rich pseudo-repeat 1"

FT Region 152..194 /note= "cysteine-rich pseudo-repeat 1"

FT Region /note= "cysteine-rich pseudo-repeat 1"

XX WO9904001-A1.

XX 28-JAN-1999.

XX 21-JUL-1998; 98WO-US15072.

XX 21-JUL-1997; 97US-0053203.

XX (ZYMO) ZYMOGENETICS INC.

XX Farrah TM;

XX WPI: 1999-132245/11.

XX N-PSDB; X07226.

PT Novel tumour necrosis factor receptor ZTNFR5 - useful for
 PT regulating maturation of TNF-ligand bearing cells
 XX
 PS Claim 1; Page 84-85; 109pp; English.
 XX
 CC This polypeptide comprises a new, secreted tumour necrosis factor
 CC receptor (see W97749), designated ZTNFR-5. Novel ZTNFR-5 encoding
 CC polynucleotides and polypeptides were initially identified by
 CC querying an expressed sequence tag (EST) database for sequences
 CC homologous to conserved motifs within the TNF receptor family.
 CC Based on this search, a contig of 16 ESTs (see X07226) was
 CC constructed. ZTNFR-5 polypeptides comprise 4 cysteine-rich repeats
 CC (see also W9750-55) that are homologous to other TNF receptors, in
 CC particular the soluble, secreted TNF receptor osteoprotegerin.
 CC ZTNFR-5 polypeptide can be prepared by recombinant methods. The
 CC polypeptide, especially the extracellular domain, can be used to
 CC generate a soluble variant of ZTNFR-5. The polypeptides and
 CC nucleic acids can be used to screen for ligands, agonists and
 CC antagonists of ZTNFR-5. The polypeptides can be used in bone cell
 CC regulation and to regulate the maturation of TNF ligand-bearing
 CC cells such as T- or B-cells, lymphocytes, peripheral blood
 CC mononuclear cells, polymorphonuclear leukocytes, fibroblasts or
 CC haematopoietic cells.
 XX
 SQ Sequence 300 AA;

Query Match 100.0%; Score 148; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 4.4e-10;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FVQPCRRDSPPTCGPCPPRHVTQ 24
 |||||
 DB 57 fvqpcrrdspctcpcprhytq 80

RESULT 12
 W95082
 ID W95082 standard; Protein; 300 AA.

XX W95082;

XX 20-MAY-1999 (first entry)

XX Orphan receptor (HUMAN NTR-1) polypeptide.

KW HUMAN NTR-1; orphan receptor; osteoprotegerin; OPG; TNFR; human;
 KW tumour necrosis factor receptor; muscle disorder; bone mass; screening;
 KW muscle metabolism; binding agent; cognate ligand.

XX Homo sapiens.

XX W09907738-A2.

XX 18-FEB-1999.

XX 04-AUG-1998; 98WO-US16202.

XX 06-AUG-1997; 97US-0054869.

XX (PROC) PROCTER & GAMBLE CO.
 XX (REGE-) REGENERON PHARM INC.

XX Maslakowski PJ, Morris J, Valenzuela DM;

XX WPI; 1999-167365/14.

XX N-PSDB; X22300.

XX Novel orphan human receptor polypeptide and nucleic acid - useful as
 XX diagnostic reagents and for treatment of muscle disorders

XX Claim 7; Page 21; 23pp; English.

XX

CC This represents a HUMAN NTR-1 polypeptide, a novel orphan receptor. The
 CC protein is related to osteoprotegerin (OPG) and to tumour necrosis factor
 CC receptor (TNFR). Host cells transformed with a vector comprising the
 CC HUMAN NTR-1 nucleic acid are used for the recombinant expression of the
 CC protein. HUMAN NTR-1 proteins and antibodies immuno specific for the
 CC protein are useful for diagnosis and treatment of humans and animals,
 CC especially muscle disorders, as the receptor is involved in regulation of
 CC bone mass and muscle metabolism. HUMAN NTR-1 receptors are also useful
 CC for screening for novel binding agents, and cognate ligands, which may be
 CC used to treat disorders associated with HUMAN NTR-1 imbalance.

XX Sequence 300 AA;

Query Match 100.0%; Score 148; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 4.4e-10;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FVQPCRRDSPPTCGPCPPRHVTQ 24
 |||||
 DB 57 fvqpcrrdspctcpcprhytq 80

RESULT 13
 Y42183

ID Y42183 standard; Protein; 302 AA.

XX Y42183;

XX 17-DEC-1999 (first entry)

XX Human FLINT #2 protein sequence.

KW Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
 KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
 KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
 KW reperfusion-associated injury; aplastic anaemia; differentiation;
 KW growth; myelodysplastic syndrome; pancytopenic condition;
 KW myocardial ischaemia.

XX Homo sapiens.

XX W09950413-A2.

XX 07-OCT-1999.

XX 30-MAR-1999; 99WO-US06797.

XX 30-MAR-1998; 98US-0079856.

XX 20-MAY-1998; 98US-0086074.

XX 09-SEP-1998; 98US-0099643.

XX 17-DEC-1998; 98US-0112577.

XX 18-DEC-1998; 98US-0112703.

XX 18-DEC-1998; 98US-0112933.

XX 22-DEC-1998; 98US-0113407.

XX (ELIL) LILLY & CO ELI.

XX Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;

XX Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;

XX Song HY, Wang J, Wu X, Zuckerman SH;

XX WPI; 1999-591319/50.

XX N-PSDB; Z25376.

XX Use of mature FLINT for treating acute liver failure, inflammation,
 XX cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
 XX and proinflammatory activity

XX Example 2; Fig 2; 99pp; English.

XX The present invention describes therapeutic applications of mature FLINT
 CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT

CC (mFLINT), which is a member of the tumour necrosis factor receptor
 CC superfamily, is used for treating acute liver failure, inflammation of
 CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
 CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
 CC injury or disorder such as hypercoagulation (including use with
 CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
 CC or disorder, Type I diabetes, cancer, cell damage or damage to an
 CC innocent bystander tissue that is induced by a chemotherapeutic agent or
 CC therapeutic irradiation, treating haematopoietic progenitor cells that
 CC have been exposed to therapeutic radiation or chemotherapy, aplastic
 CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
 CC also used for promoting the growth or differentiation of a haematopoietic
 CC progenitor cell or CD34⁺ cell and preventing damage to a cardiac myocyte
 CC resulting from abnormal myocardial ischaemia. The present sequence
 CC represents human FLINT.
 XX
 SQ Sequence 302 AA;

Query Match 100.0%; Score 148; DB 20; Length 302;
 Best Local Similarity 100.0%; Pred. No. 4.4e-10;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVQPCRRDSTTCGCPPPRHYTQ 24
 |||||
 Db 57 fvgprcdsdpmctcgcpcprhytq 80

RESULT 14
 Y77458

ID Y77458 standard; Protein: 300 AA.

XX Y77458;

DT 05-JUN-2000 (first entry)

DE Human TNF receptor-like protein, HDTEA84.

XX TNF receptor family; tumour necrosis factor; HDTEA84; HSLJD37R;

KW Rank-like protein; RANKL; Immune disorder: Inflammation; allergy;
 KW immunosuppressant; antiarthritic; antirheumatoid; antiinflammatory;
 KW dermatological; antithyroid.

XX Homo sapiens.

XX WO200001817-A2.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-US12366.

XX 06-JUL-1998; 98US-0110938.

PR 13-JUL-1998; 98US-0114466.

PR 23-JUL-1998; 98US-0093897.

PR 12-AUG-1998; 98US-0132968.

PR 18-AUG-1998; 98US-0136214.

PR 11-SEP-1998; 98US-0099999.

XX (SCHE) SCHERING CORP.

XX Bates EEM, Lebecque SJE, Murphy EE, Mattson JD, Gorman DM;

PI Hedrick JA, Wang L, Zlotnik A, Murgolo NJ, Greene JR, Johnston JA;

PI Bazan JF, Mahony D, Lees EM;

XX WPI; 2000-171015/15.

DR N-PSDB; 292404.

XX New isolated mammalian genes, used to develop products for treating

PT e.g. immune, inflammatory or allergic abnormalities, cancers or

PT degenerative conditions

XX Claim 24: Page 157; 218pp; English.

CC The invention relates to a number of primate and/or rodent proteins, and
 CC the genes which encode them. The invention encompasses human dendritic
 CC cell prostaglandin transporter (DC-PGT); the TNF (tumour necrosis
 CC factor) receptor family-related proteins HDTEA84, HSLJD37R and RANKL;
 CC human CC chemokine HCC5; human deubiquitinating proteins DUB1 and DUB
 CC 12; human MD-1 and human and murine MD-2 proteins, which exhibit the
 CC properties of ligands for proteins comprising a leucine-rich motif (LRR);
 CC human cyclin E2; cDNAs encoding these proteins; and antibodies against
 CC these proteins. The proteins can be used for modulating the physiology or
 CC development of a cell. They can be used to modulate uptake of substrates
 CC (e.g., prostaglandin-like molecules), to modulate or mediate cellular
 CC interactions (e.g., induce or prevent trafficking, proliferation, or
 CC differentiation of cells), or are intracellular proteins which are
 CC important in various cellular processes such as the deubiquitination of
 CC proteins or cell cycle regulation. The products can be used for
 CC treating medical conditions such as immune, inflammatory or allergic
 CC disorders, or abnormal cellular proliferation, for example, cancers or
 CC degenerative conditions. They can be used to modulate immune responses in
 CC disease states e.g., autoimmune disorders, including rheumatoid
 CC arthritis, systemic lupus erythematosus, Hashimoto's autoimmune
 CC thyroiditis, as well as acute and chronic inflammatory responses in which
 CC T cell activation, expansion, and/or immunological T cell memory play an
 CC important role. Sequences Y77458-Y77461 and Y77465-Y77468 represent
 CC TNF receptor family-related proteins. Y77458 is the human protein
 CC HDTEA84, Y77459-Y77461 are human HSLJD37R proteins, Y77465 is
 CC murine Rank-like protein RANKL, and Y77466-Y77468 are human RANKL
 CC proteins.
 XX
 SQ Sequence 300 AA;

Query Match 95.9%; Score 142; DB 21; Length 300;
 Best Local Similarity 95.8%; Pred. No. 2.1e-09;
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FVQPCRRDSTTCGCPPPRHYTQ 24
 |||||
 Db 57 fvgprcdsdpmctcgcpcprhytq 80

RESULT 15

R99951

ID R99951 standard; Protein: 84 AA.

XX R99951;

XX 23-APR-1997 (first entry)

XX Mutated OCIF, OCIF-CPst.

DE Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;

XX osteoporosis.

OS Synthetic.

XX Key Location/Qualifiers

FT Peptide 1..21 /note= "Signal peptide"

FT Protein 22..84 /note= "Mature OCIF-CPst"

FT WO9626217-A1.

XX 29-AUG-1996.

XX 20-FEB-1996; 96WO-JP00374.

XX 21-JUL-1995; 95JP-0207508.

PR 20-FEB-1995; 95JP-0054977.

XX (SNOW) SNOW BRAND MILK PROD CO LTD.

XX Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;

PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
XX WPI; 1996-402320/40.
DR N-PSDB; T33181.
XX
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
XX
PS Claim 89; Page 131; 183pp; Japanese.
XX
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-cpst in which amino acids 62-380 of the mature
CC OCIF protein are replaced by Leu-Val. These changes are caused by
CC the introduction of a restriction site in the DNA encoding this protein.
CC The OCIF of the invention has a molecular weight by SDS-PAGE of 60 kD.
CC under reducing conditions and 120 kD under non-reducing conditions. The
CC protein is adsorbed onto cation-exchangers or heparin and its activity is
CC lowered after 10 mins at 70 deg.C or 30 mins at 56 deg.C, and is lost
CC after 10 mins at 90 deg.C. OCIF is useful in the control of bone
CC resorption and therefore in the treatment and prevention of disorders
CC of bone resorption, e.g. osteoporosis.
XX
SQ Sequence 84 AA;

Query Match 37.8%; Score 56; DB 17; Length 84;
Best Local Similarity 34.8%; Pred. No. 5.2;
Matches 8; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 FVQPCRRDSPTTCGCPPRHYT 23
::: | | | | |
Db 49 ylkqhtakwktvcapcdhyt 71

Search completed: January 30, 2001, 16:45:29
Job time: 589 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run On: January 30, 2001, 16:50:34 ; Search time 149.64 seconds
(without alignments)
10.890 Million cell updates/sec

Title: US-09-518-931-4_COPY_57_80
Perfect score: 148
Sequence: 1 FVQPCRRDSPTTCGCPKPRHTQ 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues
Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_66: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	56	37.8	302	A31921	collagen dpy-13 pr
2	55	37.2	289	T26812	hypothetical prote
3	53.5	36.1	813	T02672	hypothetical prote
4	53	35.8	247	I48699	outer dense fiber
5	53	35.8	299	T19564	hypothetical prote
6	53	35.8	299	T29956	hypothetical prote
7	53	35.8	299	T25407	hypothetical prote
8	52.5	35.5	2318	S45306	notch 3 protein -
9	51	34.5	266	T22706	hypothetical prote
10	51	34.5	299	T22705	hypothetical prote
11	51	34.5	299	T24833	hypothetical prote
12	51	34.5	459	I48854	gene murine tumour
13	51	34.5	474	T38634	tumor necrosis fac
14	51	34.5	499	S36531	L1 protein - human
15	51	34.5	1182	I48378	hairless protein -
16	50.5	34.1	102	A37413	hypothetical RTVL-
17	50.5	34.1	501	PIWLEP	L1 protein - Europ
18	50	33.8	103	A42523	A53R protein - vac
19	50	33.8	103	J01791	Salp16R protein -
20	50	33.8	304	T22602	hypothetical prote
21	50	33.8	305	T20906	hypothetical prote
22	50	33.8	316	T20497	hypothetical prote
23	50	33.8	327	T34203	hypothetical prote
24	50	33.8	334	A48151	sperm tail protein
25	50	33.8	418	1 WZBEF9	gene 68 protein -
26	50	33.8	991	T78843	receptor protein-t
27	50	33.8	1234	1 NBMSH	complement factor
28	49.5	33.4	186	T22278	hypothetical prote
29	49.5	33.4	190	T22279	hypothetical prote

30	49.5	33.4	1299	2	T43251	furin (EC 3.4.21.7
31	49	33.1	131	1	KRG73M	keratin high-sulfu
32	49	33.1	230	2	T43487	hypothetical prote
33	49	33.1	277	2	I37552	OX40 homolog - hum
34	49	33.1	296	2	A31219	collagen 1 - Caeno
35	49	33.1	300	2	T24482	hypothetical prote
36	49	33.1	301	2	T21314	hypothetical prote
37	49	33.1	303	2	T28999	hypothetical prote
38	49	33.1	315	2	D48560	immediate-early pr
39	49	33.1	452	2	T36042	probable plasmid r
40	49	33.1	647	2	T43952	hypothetical prote
41	49	33.1	759	2	T44142	DRI protein (impor
42	49	33.1	775	1	EDBE11	immediate-early pr
43	49	33.1	896	2	B43817	transforming prote
44	49	33.1	906	2	A43817	transforming prote
45	48.5	32.8	1663	2	T42092	s-afadin - rat

ALIGNMENTS

RESULT 1
A31921
collagen dpy-13 precursor - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 21-May-1990 #sequence_revision 21-May-1990 #text_change 29-Oct-1999
C:Accession: A31921; T29577
R:Yvon Mende, N.; Bird, D.M.; Albert, P.S.; Riddle, D.L.
Cell 55, 567-576, 1988
A:Title: dpy-13: a nematode collagen gene that affects body shape.
A:Reference number: A31921; MUID:89028668
A:Accession: A31921
A:Molecule type: DNA
A:Residues: 1-302 <VON>
A:Cross-references: GB:M23559; NID:gl56269; PIDN:AAA27994.1; PID:gl56270
R:Miller, N.; Bradshaw, H.
Submitted to the EMBL Data Library, December 1995
A:Description: The sequence of C. elegans cosmid F30B5.
A:Reference number: Z20644
A:Accession: T29577
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: DNA
A:Residues: 1-302 <MIL>
A:Cross-references: EMBL:U42437; PIDN:AAA83499.1; CESP:F30B5.1
C:Genetics:
A:Gene: CESP:F30B5.1
A:Introns: 27/3; 246/1
C:Superfamily: unassigned collagens

Query Match 37.8%; Score 56; DB 2; Length 302;
Best Local Similarity 50.0%; Pred. No. 5.8;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QRPCRRDSPTTCGCP 18
DB 137 QRPCERTPPCKPCP 152

RESULT 2
T26812
hypothetical protein Y41E3.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
R:McMurray, A.
Submitted to the EMBL Data Library, May 1997
A:Reference number: Z20270
A:Accession: T26812
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: DNA
A:Residues: 1-289 <WIL>
A:Cross-references: EMBL:Z95559; PIDN:CAB08999.1; GSPDB:GN00022; CESP:Y41E3.2

A: Experimental source: clone Y41E3
C: Genetics:
A: Gene: CESP:Y41E3.2
A: Map position: 4
A: Introns: 27/3; 273/2
C: Superfamily: unassigned collagens

Query Match 37.2%; Score 55; DB 2; Length 289;
Best Local Similarity 50.0%; Pred. No. 7.4;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QRPCRRDSPTTCGCP 18
|:|:| |:| |:|

Db 124 QQPCPTTPPPCPCP 139

RESULT 3

T02672
hypothetical protein R31449_3 - human (fragment)
C: Species: Homo sapiens (man)
C: Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
C: Accession: T02672
R: Lamerdin, J. E.; McCready, P. M.; Skowronski, E.; Adamson, A. W.; Burkhart-Schultz, K.; C
J.; Danganan, L.; Poundstone, P.; Christensen, M.; Georgescu, A.; Avila, J.; Liu, S.; A
P.; Quan, G.; Krommiller, B.; Arellano, A.; Montgomery, M.; Ow, D.; Nolan, M.;
submitted to the EMBL Data Library, June 1998
A: Authors: Trong, S.; Kobayashi, A.; Olsen, A. S.; Carrano, A. V.
A: Description: Sequence analysis of a 3.5 Mb contig in human 19p13.3 containing a serine
A: Reference number: Z14696
A: Accession: T02672
A: Status: preliminary; translated from GB/EMBL/DDBJ
A: Molecule type: DNA
A: Residues: 1-813 <LAM>
A: Cross-references: EMBL:AC005175; NID:G3253116; PID:G3253120
C: Genetics:
A: Map position: 19
A: Introns: 93/2; 125/2; 174/1; 228/2; 266/3; 331/1; 372/1; 474/3; 637/1; 678/3
A: Note: R31449_3

Query Match 36.1%; Score 53.5; DB 2; Length 813;
Best Local Similarity 52.4%; Pred. No. 25;
Matches 11; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

QY 4 RPCRDR--SPTTCGCPPRH 21
|:|:| |:| |:|

Db 645 RHCRRDPLWPTLCRDWPPH 665

RESULT 4

I48699
outer dense fiber protein - mouse
C: Species: Mus musculus (house mouse)
C: Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C: Accession: I48699; S51743
R: Hoyer-Fender, S.; Burfeind, P.; Hameister, H.
Cytogenet. Cell Genet. 70, 200-204, 1995
A: Title: Sequence of mouse Odf1 cDNA and its chromosomal localization: extension of the
A: Reference number: I48699; MUID:95309020
A: Accession: I48699
A: Status: preliminary; translated from GB/EMBL/DDBJ
A: Molecule type: mRNA
A: Residues: 1-247 <RES>
A: Cross-references: EMBL:X79446; NID:G607139; PIDN:CAA55965.1; PID:G607140
R: Hoyer-Fender, S.; Burfeind, P.; Hameister, H.
submitted to the EMBL Data Library, May 1994
A: Description: Sequence of mouse ODF1 cDNA and its chromosomal localization: extension of
A: Reference number: S51743
A: Accession: S51743
A: Status: preliminary
A: Molecule type: mRNA
A: Residues: 1-247 <HOY>

A: Cross-references: EMBL:X79446; NID:G607139; PIDN:CAA55965.1; PID:G607140

Query Match 35.8%; Score 53; DB 2; Length 247;
Best Local Similarity 53.3%; Pred. No. 11;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 5 PCRRDSPTTCGCP 19

|:| |:| |:|

Db 203 PCSPCAPCAGPCGP 217

RESULT 5

T19564
hypothetical protein C29F4.1 - Caenorhabditis elegans
C: Species: Caenorhabditis elegans
C: Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C: Accession: T19564
R: Kershaw, J.
submitted to the EMBL Data Library, January 1996
A: Reference number: Z19143
A: Accession: T19564
A: Status: preliminary; translated from GB/EMBL/DDBJ
A: Molecule type: DNA
A: Residues: 1-299 <MIL>
A: Cross-references: EMBL:Z68335; PIDN:CAA92729.1; GSPDB:GN00022; CESP:C29F4.1
A: Experimental source: clone C29F4
C: Genetics:
A: Gene: CESP:C29F4.1
A: Map position: 4
A: Introns: 27/3
C: Superfamily: unassigned collagens

Query Match 35.8%; Score 53; DB 2; Length 299;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QRPCRRDSPTTCGCP 18

|:|:| |:| |:|

Db 134 QQPCPTTPPPCKPCP 149

RESULT 6

T29956
hypothetical protein F36A4.10 - Caenorhabditis elegans
C: Species: Caenorhabditis elegans
C: Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C: Accession: T29956
R: Miller, N.; Bradshaw, H.
submitted to the EMBL Data Library, April 1996
A: Description: The sequence of C. elegans cosmid F36A4.
A: Reference number: Z20713
A: Accession: T29956
A: Status: preliminary; translated from GB/EMBL/DDBJ
A: Molecule type: DNA
A: Residues: 1-299 <MIL>
A: Cross-references: EMBL:U53333; PIDN:AAA96155.1; GSPDB:GN00022; CESP:F36A4.10
A: Experimental source: strain Bristol N2; clone F36A4
C: Genetics:
A: Gene: CESP:F36A4.10
A: Map position: 4
A: Introns: 27/3; 243/1
C: Superfamily: unassigned collagens

Query Match 35.8%; Score 53; DB 2; Length 299;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QRPCRRDSPTTCGCP 18

|:|:| |:| |:|

Db 134 QQPCPTTPPPCKPCP 149

```
RESULT 7
T25407
hypothetical protein T28C6.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T25407; T25409
R:Lloyd, C.
submitted to the EMBL Data Library, September 1995
A:Reference number: T20030
A:Accession: T25407
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-299 <WIL>
A:Cross-references: EMBL:Z54238; PIDN:CAA90995.1; GSPDB:GN00022; CESP:T28C6.4
A:Experimental source: clone T28C6
A:Accession: T25409
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-299 <W12>
A:Cross-references: EMBL:Z54238; PIDN:CAA90997.1; GSPDB:GN00022; CESP:T28C6.6
A:Experimental source: clone T28C6
C:Genetics:
A:Gene: CESP:T28C6.4; CESP:T28C6.6
A:Map position: 4
A:Introns: 27/3
C:Superfamily: unassigned collagens

Query Match 35.8%; Score 53; DB 2; Length 299;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 3 QRPCRRDSPTTCGCP 18
I:| | :| | | |
Db 134 QQPCDPITPPCKPCP 149

RESULT 8
S45306
notch 3 protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Sep-1999
C:Accession: S45306
R:Lardelli, M.; Dahlstrand, J.; Lendahl, U.
Mech. Dev. 46, 123-136, 1994
A:Title: The novel Notch homologue mouse Notch 3 lacks specific epidermal growth factor-
A:Reference number: S45306; MUID:95001556
A:Accession: S45306
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2318 <LAR>
A:Cross-references: EMBL:X74760; NID:9483580; PIDN:CAA52776.1; PID:9483581
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
F:163-195/Domain: EGF homology <EGF1>
F:474-505/Domain: EGF homology <EGF>
F:854-885/Domain: EGF homology <EGF2>
F:1839-1871/Domain: ankyrin repeat homology <AN1>
F:1872-1904/Domain: ankyrin repeat homology <AN2>
F:1906-1938/Domain: ankyrin repeat homology <AN3>
F:1939-1971/Domain: ankyrin repeat homology <AN4>
F:1972-2004/Domain: ankyrin repeat homology <AN5>

Query Match 35.5%; Score 52.5; DB 2; Length 2318;
Best Local Similarity 36.7%; Pred. No. 76;
Matches 11; Conservative 3; Mismatches 9; Indels 7; Gaps 1;

Qy 2 VORPCRRDSPTTCGP-----CPRHYTQ 24
I:| | | :| | |
Db 125 VSRPCVHGAPCVSGPDGRFACACPPGQ 154

RESULT 9
T22706
hypothetical protein F55C10.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T22706
R:Dobson, R.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z19603
A:Accession: T22706
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-266 <WIL>
A:Cross-references: EMBL:Z74036; PIDN:CAA98487.1; GSPDB:GN00023; CESP:F55C10.3
A:Experimental source: clone F55C10
C:Genetics:
A:Gene: CESP:F55C10.3
A:Map position: 5
A:Introns: 21/3
C:Superfamily: unassigned collagens

Query Match 34.5%; Score 51; DB 2; Length 266;
Best Local Similarity 50.0%; Pred. No. 21;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 3 QRPCRRDSPTTCGCP 18
I:| | :| | | |
Db 101 QQPCDPITPPCKPCP 116

RESULT 10
T22705
hypothetical protein F55C10.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T22705
R:Dobson, R.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z19603
A:Accession: T22705
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-299 <WIL>
A:Cross-references: EMBL:Z74036; PIDN:CAA98486.1; GSPDB:GN00023; CESP:F55C10.2
A:Experimental source: clone F55C10
C:Genetics:
A:Gene: CESP:F55C10.2
A:Map position: 5
A:Introns: 54/3
C:Superfamily: unassigned collagens

Query Match 34.5%; Score 51; DB 2; Length 299;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 3 QRPCRRDSPTTCGCP 18
I:| | :| | | |
Db 134 QQPCDPITPPCKPCP 149

RESULT 11
T24833
hypothetical protein T11F9.9 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T24833
R:Lennard, N.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z19941
A:Accession: T24833
```

A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-299 <WIL>
A:Cross-references: EMBL:274042; PIDN:CAA98525.1; GSPDB:GN00023; CESP:Tllf9.9
A:Experimental source: clone Tllf9
C:Genetics:
A:Gene: CESP:Tllf9.9
A:Map position: 5
A:Introns: 54/3
C:Superfamily: unassigned collagens

Query Match 34.5%; Score 51; DB 2; Length 299;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QRCRRDSDPTTCGCPCP 18
: : : : :
Db 134 QPCDPTTPPCQCP 149

RESULT 12
I48854
gene murine tumor necrosis factor receptor 2 protein - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 23-Jul-1999
C:Accession: I48854
R:Powell, E.E.; Wicker, L.S.; Peterson, L.B.; Todd, J.A.
Mamm. Genome 5, 726-727, 1994
A:Title: Allelic variation of the type 2 tumor necrosis factor receptor gene.
A:Reference number: I48854; MUID:95178848
A:Accession: I48854
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-459 <RES>
A:Cross-references: EMBL:X76401; NID:g433830; PIDN:CAA53981.1; PID:g433831
C:Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology
F:151-188/Domain: NGF receptor repeat homology <NGF>

Query Match 34.5%; Score 51; DB 2; Length 459;
Best Local Similarity 37.3%; Pred. No. 33;
Matches 9; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 1 FVQRPCCRDSPTTCGCPPPRHVTQ 24
: : : : :
Db 48 YVKHFCNKTSDTVACDEASMYTQ 71

RESULT 13
B38634
tumor necrosis factor receptor type 2 precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
A:Accession: B38634; A40254; S54816
R:Lewis, M.; Tartaglia, L.A.; Lee, A.; Bennett, G.L.; Rice, G.C.; Wong, G.H.W.; Chen, E.
Proc. Natl. Acad. Sci. U.S.A. 88, 2830-2834, 1991
A:Title: Cloning and expression of cDNAs for two distinct murine tumor necrosis factor
A:Reference number: B38634; MUID:91187885
A:Accession: B38634
A:Molecule type: mRNA
A:Residues: 1-474 <LEW>
A:Cross-references: GB:M60469; NID:g199827; PIDN:AAA39752.1; PID:g199828
R:Goodwin, R.G.; Anderson, D.; Jerzy, R.; Davis, T.; Brannan, C.I.; Copeland, N.G.; Jenk
Mol. Cell. Biol. 11, 3020-3026, 1991
A:Title: Molecular cloning and expression of the type 1 and type 2 murine receptors for
A:Reference number: A40254; MUID:91246168
A:Accession: A40254
A:Molecule type: mRNA
A:Residues: 1-474 <GOO>
A:Cross-references: GB:M60469; NID:g199827; PIDN:AAA39752.1; PID:g199828
R:Kisomergis, M.; Fellows, R.; Feldmann, M.; Chernajovsky, Y.
submitted to the EMBL Data Library, May 1995

A:Description: Characterization of the promoter region of the murine p75-TNF receptor
A:Reference number: S54816
A:Accession: S54816
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-22 <KIS>
A:Cross-references: EMBL:X87128; NID:g809043; PIDN:CAA60618.1; PID:g809044
C:Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology
C:Keywords: cytokine receptor; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-474/Product: tumor necrosis factor receptor type 2 #status predicted <MAT>
F:40-77/Domain: NGF receptor repeat homology <NG1>
F:79-120/Domain: NGF receptor repeat homology <NG2>
F:166-203/Domain: NGF receptor repeat homology <NG4>

Query Match 34.5%; Score 51; DB 2; Length 474;
Best Local Similarity 37.5%; Pred. No. 33;
Matches 9; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 1 FVQRPCCRDSPTTCGCPPPRHVTQ 24
: : : : :
Db 63 YVKHFCNKTSDTVACDEASMYTQ 86

RESULT 14
S36531
L1 protein - human papillomavirus type 53
C:Species: human papillomavirus type 53
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S36531
R:Deilus, H.; Hofmann, B.
submitted to the EMBL Data Library, August 1993
A:Description: Primer-directed sequencing of human papillomavirus types.
A:Reference number: S36469
A:Accession: S36531
A:Molecule type: DNA
A:Residues: 1-499
A:Cross-references: EMBL:X74482; NID:g997046; PIDN:CAA52595.1; PID:g997052
C:Superfamily: papillomavirus L1 protein
C:Keywords: late protein

Query Match 34.5%; Score 51; DB 2; Length 499;
Best Local Similarity 66.7%; Pred. No. 35;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 8 RDSPTTCGCPCP 19
: : : : :
Db 174 RSTPTTAGDCPP 185

RESULT 15
I48378
hairless protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C:Accession: I48378
R:Cachon-Gonzalez, M.B.; Fenner, S.; Coffin, J.M.; Moran, C.; Stoye, J.P.
Proc. Natl. Acad. Sci. U.S.A. 91, 7717-7721, 1994
A:Title: Structure and expression of the hairless gene of mice.
A:Reference number: I48378; MUID:94329587
A:Accession: I48378
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-1182 <RES>
A:Cross-references: EMBL:232675; NID:g531706; PIDN:CAA83587.1; PID:g531707

Query Match 34.5%; Score 51; DB 2; Length 1182;
Best Local Similarity 34.2%; Pred. No. 68;
Matches 13; Conservative 4; Mismatches 3; Indels 18; Gaps 3;

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 17:06:08 ; Search time 80.04 Seconds
(without alignments)
9.683 Million cell updates/sec

Title: US-09-518-931-4_COPY_57_80
Perfect score: 148
Sequence: 1 FVQPCGRDSPTTCGCPRRHYTQ 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	56	37.8	302	1 CDC_CAEEL	P17657 caenorhabdi
2	53	35.8	247	1 ODFP_MOUSE	Q61999 mus musculu
3	52.5	35.5	2318	1 NYC3_MOUSE	Q61982 mus musculu
4	51	34.5	266	1 YXWK_CAEEL	Q21184 caenorhabdi
5	51	34.5	474	1 TNR2_MOUSE	P25119 mus musculu
6	51	34.5	499	1 VIL1_HPV53	Q05113 human papil
7	51	34.5	1182	1 HAIR_MOUSE	Q61645 mus musculu
8	50.5	34.1	501	1 VIL1_PAPVE	P11326 european el
9	50	33.8	103	1 VA53_VACCC	P21071 vaccinia vi
10	50	33.8	103	1 VA53_VACCV	P24756 vaccinia vi
11	50	33.8	418	1 US02_HSVB	P28964 equine herp
12	50	33.8	1037	1 EPA5_HUMAN	P54756 homo sapien
13	50	33.8	1234	1 CFAH_MOUSE	P06909 mus musculu
14	49.5	33.4	561	1 EFS_HUMAN	O43281 homo sapien
15	49	33.1	277	1 OX40_HUMAN	P43489 homo sapien
16	49	33.1	296	1 CC01_CAEEL	P08124 caenorhabdi
17	49	33.1	415	1 TNRC_MOUSE	P50284 mus musculu
18	49	33.1	775	1 ICP0_HSV11	P08393 herpes simp
19	49	33.1	896	1 CBL_MOUSE	P22682 mus musculu
20	49	33.1	906	1 CBL_HUMAN	P22681 homo sapien
21	48	32.4	271	1 OX40_RAT	P15725 rattus norv
22	48	32.4	458	1 HK51_MOUSE	P42581 mus musculu
23	48	32.4	983	1 EPA3_MOUSE	P29319 mus musculu
24	48	32.4	1005	1 EPA5_RAT	P54757 rattus norv
25	48	32.4	1196	1 DNB1_HSV2H	P89452 herpes simp
26	48	32.4	1197	1 DNB1_HSV2	P36384 herpes simp
27	48	32.4	1446	1 IE18_PRRVA	P33479 pseudorabie
28	48	32.4	1839	1 POLR_EPMV	P20126 eggplant mo
29	47.5	32.1	116	1 UL81_HCMVA	P16825 human cytom
30	47.5	32.1	245	1 ODFP_RAT	P21769 rattus norv
31	47.5	32.1	361	1 COOH_RHORU	P31895 rhodospiril
32	47	31.8	131	1 KRA3_SHEEP	P02441 ovilis aries
33	47	31.8	132	1 KRA3_CAPHI	P02442 capra hircu

34	47	31.8	272	1 OX40_MOUSE	P47741 mus musculu
35	47	31.8	903	1 ECLC_BOVIN	P54281 bos taurus
36	46.5	31.4	130	1 KR3A_SHEEP	P02443 ovilis aries
37	46.5	31.4	1104	1 NFXL_HUMAN	Q12986 homo sapien
38	46.5	31.4	2813	1 VNF_HUMAN	P04275 homo sapien
39	46	31.1	68	1 M84D_DROME	Q01645 drosophila
40	46	31.1	414	1 LMP1_CHICK	P05300 gallus gall
41	46	31.1	568	1 MASY_CUCSA	P08216 cecumis sat
42	46	31.1	822	1 FGR1_RAT	Q04589 rattus norv
43	46	31.1	2813	1 VNF_CANFA	Q28295 canis famil
44	45.5	30.7	152	1 CORA_RAT	Q63532 rattus norv
45	45.5	30.7	513	1 VIL1_PAPVD	P03104 deer papill

ALIGNMENTS

RESULT 1
CCDC_CAEEL STANDARD; PRT; 302 AA.
AC P17657;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE CUTICLE COLLAGEN DPY-13.
GN DPY-13 OR DPY-16 OR F30B5.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89028668; PubMed=2846184;
RA von Mende N., Bird D.M., Albert P.S., Riddle D.L.;
RT "dpy-13: a nematode collagen gene that affects body shape.";
RL Cell 55:567-576(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP VARIANT DPY13(E225).
RX MEDLINE=93013043; PubMed=1398138;
RA Bird D.M.;
RT "Sequence comparison of the Caenorhabditis elegans dpy-13 and col-34 genes, and their deduced collagen products.";
RL Gene 120:261-266(1992).
CC -!- FUNCTION: NEMATODE CUTICLES ARE COMPOSED LARGELY OF COLLAGEN-LIKE PROTEINS. THE CUTICLE FUNCTIONS BOTH AS AN EXOSKELETON AND AS A BARRIER TO PROTECT THE WORM FROM ITS ENVIRONMENT.
CC -!- SUBUNIT: COLLAGEN POLYPEPTIDE CHAINS ARE COMPLEXED WITHIN THE CUTICLE BY DISULFIDE BONDS AND OTHER TYPES OF COVALENT CROSS-LINKS.
CC -!- DISEASE: MUTATIONS IN DPY-13 AFFECTS THE BODY SHAPE.
CC -!- SIMILARITY: TO OTHER COLLAGENS. STRONG, TO OTHER CUTICLE COLLAGENS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; M23559; AAA27994.1; .
CC EMBL; U42437; AAA83499.1; .
DR PIR; A31921; A31921.
DR WORMPEP; F30B5.1; CE04462.
DR INTERPRO; IPR000087; .
DR PFW; PF01391; Collagen; 2.
KW Cuticle; Connective tissue; Repeat; Multigene family; Collagen.
FT DOMAIN 106 135
TRIPLE-HELICAL REGION.

```

FT DOMAIN 154 210 TRIPLE-HELICAL REGION.
FT DOMAIN 219 278 TRIPLE-HELICAL REGION.
FT VARIANT 205 205 G -> E (IN DRY13(E225)).
SQ SEQUENCE 302 AA; 30052 MW; 16B5CDAF14E0F0E5 CRC64;

Query Match 37.8%; Score 56; DB 1; Length 302;
Best Local Similarity 50.0%; Pred. No. 1.8;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QRCRDSPPTTCGPGP 18
   ||| :| |||
Db 137 QKCEETTPPCPGP 152

RESULT 2
ID ODFP_MOUSE STANDARD; PRT; 247 AA.
AC Q61999;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 01-NOV-1997 (Rel. 35, Last annotation update)
DE OUTER DENSE FIBER PROTEIN.
GN ODF1 OR ODFP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-TESTIS;
RC MEDLINE=95309020; PubMed=7789171;
RA Hoyer-Fender S., Burfeind P., Hameister H.;
RT "Sequence of mouse Odf1 cDNA and its chromosomal localization:
RT extension of the linkage group between human chromosome 8 and mouse
RT chromosome 15.";
RL Cytogenet. Cell Genet. 70:200-204(1995).
CC -1- FUNCTION: COMPONENT OF THE OUTER DENSE FIBERS (ODF) OF
CC SPERMATOZOEA. ODF ARE FILAMENTOUS STRUCTURES LOCATED ON THE OUTSIDE
CC OF THE AXONEME IN THE MIDPIECE AND PRINCIPAL PIECE OF THE
CC MAMMALIAN SPERM TAIL AND MAY HELP TO MAINTAIN THE PASSIVE ELASTIC
CC STRUCTURES AND ELASTIC RECOIL OF THE SPERM TAIL.
CC -1- DOMAIN: THE C-TERMINAL CONTAINS MANY C-X-P REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X79446; CAA55965.1; --
DR MGD; MGI:97424; ODF1.
KW Repeat; Spermatogenesis.
FT DOMAIN 34 78 2 X 5 AA REPEATS OF [RC]-C-L-C-D.
FT REPEAT 34 38 1.
FT REPEAT 74 78 2.
FT DOMAIN 195 235 C-X-P REPEAT REGION.
SQ SEQUENCE 247 AA; 27539 MW; 821EB57C2D67E458 CRC64;

Query Match 35.8%; Score 53; DB 1; Length 247;
Best Local Similarity 53.3%; Pred. No. 3.5;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 5 PCRRDSPPTTCGPGP 19
   || :| ||| |
Db 203 PCGPCAPCAGCGPGP 217

RESULT 3
ID NTC3_MOUSE STANDARD; PRT; 2318 AA.

```

```

AC Q61982;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE NEUROGENIC LOCUS NOTCH 3 PROTEIN.
GN NOTCH3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR X SWISS WEBSTER;
RC MEDLINE=95001556; PubMed=7918097;
RA Lardelli M., Dalstrand J., Lendahl U.;
RT "The novel Notch homologue mouse Notch 3 lacks specific epidermal
RT growth factor-repeats and is expressed in proliferating
RT neuroepithelium.";
RL Mech. Dev. 46:123-136(1994).
CC -1- FUNCTION: NOTCH 1, 2 AND 3 PLAY A COMBINATIONAL ROLE DURING
CC DEVELOPING CNS AND PROBABLY OTHER REGIONS OF THE EMBRYO.
CC -1- TISSUE SPECIFICITY: PROLIFERATING NEUROEPITHELIUM.
CC -1- DEVELOPMENTAL STAGE: CNS DEVELOPMENT.
CC -1- SIMILARITY: CONTAINS 34 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC -1- SIMILARITY: CONTAINS 6 CDC10/SWI6 REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X74760; CAA52776.1; --
DR HSP; P00740; IIXA.
DR MGD; MGI:99460; NOTCH3.
DR INTERPRO; IPR00152; --
DR INTERPRO; IPR000561; --
DR INTERPRO; IPR000800; --
DR INTERPRO; IPR001438; --
DR INTERPRO; IPR001861; --
DR INTERPRO; IPR002110; --
DR PFAM; PF00008; EGF_34.
DR PFAM; PF00023; ank_6.
DR PRINTS; PR00066; notch_3.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 18.
DR PROSITE; PS00022; EGF_1; 33.
DR PROSITE; PS01186; EGF_2; 27.
DR PROSITE; PS01187; EGF_CA; 17.
DR Differentiation; Neurogenesis; Repeat; EGF-like domain; Transmembrane;
KW Glycoprotein.
FT DOMAIN 1 1643 EXTRACELLULAR.
FT TRANSMEM 1644 1664 POTENTIAL.
FT DOMAIN 1665 2318 CYTOPLASMIC.
FT DOMAIN 39 1374 34 X EGF-TYPE REPEATS.
FT DOMAIN 1388 1503 3 X LIN/NOTCH REPEATS.
FT DOMAIN 1784 1998 6 X CDC10/SWI6 REPEATS.
FT DOMAIN 2242 2261 PEST.
FT DOMAIN 39 78 EGF-LIKE 1.
FT DOMAIN 79 119 EGF-LIKE 2.
FT DOMAIN 120 157 EGF-LIKE 3.
FT DOMAIN 159 196 EGF-LIKE 4.
FT DOMAIN 198 235 EGF-LIKE 5.
FT DOMAIN 237 273 EGF-LIKE 6.
FT DOMAIN 275 313 EGF-LIKE 7.
FT DOMAIN 315 351 EGF-LIKE 8.
FT DOMAIN 352 390 EGF-LIKE 9.
FT DOMAIN 392 430 EGF-LIKE 10, CALCIUM-BINDING (POTENTIAL).

```



```
FT DOMAIN 432 468 EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL). 555 569 FT DISULFID BY SIMILARITY.
FT DOMAIN 506 EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL). 571 580 FT DISULFID BY SIMILARITY.
FT DOMAIN 544 EGF-LIKE 13, CALCIUM-BINDING (POTENTIAL). 587 598 FT DISULFID BY SIMILARITY.
FT DOMAIN 581 EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL). 592 607 FT DISULFID BY SIMILARITY.
FT DOMAIN 593 EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL). 609 618 FT DISULFID BY SIMILARITY.
FT DOMAIN 621 EGF-LIKE 16, CALCIUM-BINDING (POTENTIAL). 625 635 FT DISULFID BY SIMILARITY.
FT DOMAIN 658 EGF-LIKE 17, CALCIUM-BINDING (POTENTIAL). 630 644 FT DISULFID BY SIMILARITY.
FT DOMAIN 694 EGF-LIKE 18, 646 655 FT DISULFID BY SIMILARITY.
FT DOMAIN 731 EGF-LIKE 19, 662 673 FT DISULFID BY SIMILARITY.
FT DOMAIN 772 EGF-LIKE 20, 687 692 FT DISULFID BY SIMILARITY.
FT DOMAIN 811 EGF-LIKE 21, CALCIUM-BINDING (POTENTIAL). 684 693 FT DISULFID BY SIMILARITY.
FT DOMAIN 850 EGF-LIKE 22, CALCIUM-BINDING (POTENTIAL). 700 710 FT DISULFID BY SIMILARITY.
FT DOMAIN 888 EGF-LIKE 23, CALCIUM-BINDING (POTENTIAL). 705 719 FT DISULFID BY SIMILARITY.
FT DOMAIN 925 EGF-LIKE 24, 721 730 FT DISULFID BY SIMILARITY.
FT DOMAIN 963 EGF-LIKE 25, 739 750 FT DISULFID BY SIMILARITY.
FT DOMAIN 1001 EGF-LIKE 26, 744 759 FT DISULFID BY SIMILARITY.
FT DOMAIN 1037 EGF-LIKE 27, 761 770 FT DISULFID BY SIMILARITY.
FT DOMAIN 1085 EGF-LIKE 28, 776 787 FT DISULFID BY SIMILARITY.
FT DOMAIN 1123 EGF-LIKE 29, 781 797 FT DISULFID BY SIMILARITY.
FT DOMAIN 1159 EGF-LIKE 30, CALCIUM-BINDING (POTENTIAL). 799 808 FT DISULFID BY SIMILARITY.
FT DOMAIN 1161 EGF-LIKE 31, 815 827 FT DISULFID BY SIMILARITY.
FT DOMAIN 1206 EGF-LIKE 32, 821 836 FT DISULFID BY SIMILARITY.
FT DOMAIN 1247 EGF-LIKE 33, 838 847 FT DISULFID BY SIMILARITY.
FT DOMAIN 1290 EGF-LIKE 34, 854 865 FT DISULFID BY SIMILARITY.
FT DOMAIN 1336 LIN/NOTCH 1, 874 885 FT DISULFID BY SIMILARITY.
FT DOMAIN 1388 LIN/NOTCH 2, 876 885 FT DISULFID BY SIMILARITY.
FT DOMAIN 1429 LIN/NOTCH 3, 892 902 FT DISULFID BY SIMILARITY.
FT DOMAIN 1467 CDC10/SWI6 1, 897 911 FT DISULFID BY SIMILARITY.
FT DOMAIN 1488 CDC10/SWI6 2, 913 922 FT DISULFID BY SIMILARITY.
FT DOMAIN 1566 CDC10/SWI6 3, 929 940 FT DISULFID BY SIMILARITY.
FT DOMAIN 1617 CDC10/SWI6 4, 949 951 FT DISULFID BY SIMILARITY.
FT DOMAIN 1661 CDC10/SWI6 5, 951 960 FT DISULFID BY SIMILARITY.
FT DOMAIN 1706 CDC10/SWI6 6, 967 978 FT DISULFID BY SIMILARITY.
FT DISULFID 43 55 BY SIMILARITY. 972 987 FT DISULFID BY SIMILARITY.
FT DISULFID 49 66 BY SIMILARITY. 989 998 FT DISULFID BY SIMILARITY.
FT DISULFID 68 77 BY SIMILARITY. 1005 1016 FT DISULFID BY SIMILARITY.
FT DISULFID 83 94 BY SIMILARITY. 1010 1023 FT DISULFID BY SIMILARITY.
FT DISULFID 88 107 BY SIMILARITY. 1025 1034 FT DISULFID BY SIMILARITY.
FT DISULFID 109 118 BY SIMILARITY. 1041 1062 FT DISULFID BY SIMILARITY.
FT DISULFID 124 135 BY SIMILARITY. 1056 1071 FT DISULFID BY SIMILARITY.
FT DISULFID 129 145 BY SIMILARITY. 1073 1082 FT DISULFID BY SIMILARITY.
FT DISULFID 147 156 BY SIMILARITY. 1089 1100 FT DISULFID BY SIMILARITY.
FT DISULFID 163 175 BY SIMILARITY. 1094 1109 FT DISULFID BY SIMILARITY.
FT DISULFID 169 184 BY SIMILARITY. 1111 1120 FT DISULFID BY SIMILARITY.
FT DISULFID 186 195 BY SIMILARITY. 1127 1138 FT DISULFID BY SIMILARITY.
FT DISULFID 202 213 BY SIMILARITY. 1132 1147 FT DISULFID BY SIMILARITY.
FT DISULFID 207 223 BY SIMILARITY. 1149 1158 FT DISULFID BY SIMILARITY.
FT DISULFID 225 234 BY SIMILARITY. 1165 1183 FT DISULFID BY SIMILARITY.
FT DISULFID 241 252 BY SIMILARITY. 1177 1192 FT DISULFID BY SIMILARITY.
FT DISULFID 246 261 BY SIMILARITY. 1194 1203 FT DISULFID BY SIMILARITY.
FT DISULFID 263 272 BY SIMILARITY. 1210 1223 FT DISULFID BY SIMILARITY.
FT DISULFID 279 292 BY SIMILARITY. 1215 1233 FT DISULFID BY SIMILARITY.
FT DISULFID 286 301 BY SIMILARITY. 1235 1244 FT DISULFID BY SIMILARITY.
FT DISULFID 303 312 BY SIMILARITY.
FT DISULFID 319 330 BY SIMILARITY.
FT DISULFID 324 339 BY SIMILARITY.
FT DISULFID 341 350 BY SIMILARITY.
FT DISULFID 356 367 BY SIMILARITY.
FT DISULFID 361 378 BY SIMILARITY.
FT DISULFID 380 389 BY SIMILARITY.
FT DISULFID 396 409 BY SIMILARITY.
FT DISULFID 403 418 BY SIMILARITY.
FT DISULFID 420 429 BY SIMILARITY.
FT DISULFID 436 447 BY SIMILARITY.
FT DISULFID 441 456 BY SIMILARITY.
FT DISULFID 458 467 BY SIMILARITY.
FT DISULFID 474 485 BY SIMILARITY.
FT DISULFID 479 494 BY SIMILARITY.
FT DISULFID 496 505 BY SIMILARITY.
FT DISULFID 512 523 BY SIMILARITY.
FT DISULFID 517 532 BY SIMILARITY.
FT DISULFID 534 543 BY SIMILARITY.
FT DISULFID 550 560 BY SIMILARITY.
```

Query Match 35.5%; Score 52.5; DB 1; Length 2318;

Best Local Similarity 36.7%; Pred. No. 29;

Matches 11; Conservative 3; Mismatches 9; Indels 7; Gaps 1;

QY 2 VQPCRRDSPTTCG-----CPPRHYTQ 24

| | | | : | : | | | : |

Db 125 VSRPCVHGAPCSVGPDPGRFACACPPGCGQ 154

RESULT 4

YXWK_CAEEL

ID YXWK_CAEEL STANDARD; PRT; 266 AA.

AC Q21184: Q20807:

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DE 01-NOV-1997 (Rel. 35, Last annotation update)

DE PUTATIVE CUTICLE COLLAGEN F55C10.3.

GN F55C10.3.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA White S.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RC STRAIN-BRISTOL N2;
RA Jones S.J.M.;
RL submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: NEMATODE CUTICLES ARE COMPOSED LARGELY OF COLLAGEN-LIKE
CC PROTEINS. THE CUTICLE FUNCTIONS BOTH AS AN EXOSKELETON AND AS A
CC BARRIER TO PROTECT THE WORM FROM ITS ENVIRONMENT (BY SIMILARITY).
CC -!- SUBUNIT: COLLAGEN POLYPEPTIDE CHAINS ARE COMPLEXED WITHIN THE
CC CUTICLE BY DISULFIDE BONDS AND OTHER TYPES OF COVALENT CROSS-
CC LINKS (BY SIMILARITY).
CC -!- SIMILARITY: TO OTHER COLLAGENS. STRONG, TO OTHER CUTICLE
CC COLLAGENS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z74036; CAA98487.1; -
DR WORMPEP; F55C10.3; CE11182.
DR INTERPRO: IPR000087; -
DR PFAM: PF01391; Collagen; 2.
KW Hypothetical protein: Cuticle; Connective tissue; Repeat;
KW Collagen; Multigene family.
FT DOMAIN 70 99 TRIPLE-HELICAL REGION.
FT DOMAIN 118 144 TRIPLE-HELICAL REGION.
FT DOMAIN 148 169 TRIPLE-HELICAL REGION.
FT DOMAIN 183 245 TRIPLE-HELICAL REGION.
SQ SEQUENCE 266 AA; 25616 MW; 984556680F1AAAF22 CRC64;

Query Match 34.5%; Score 51; DB 1; Length 266;
Best Local Similarity 50.0%; Pred. No. 6.7;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QRPCRRDSPTTCGPCP 18
Db 101 QQPCDPITPPPCQPCP 116

RESULT 5
TNR2_MOUSE
ID TNR2_MOUSE STANDARD; PRT; 474 AA.
AC P25119; P97893;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR 2 PRECURSOR (TNF-R2) (P75).
GN TNFRSF5 OR TNFR2 OR TNFR-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-91187885; PubMed-1849278;
RA Lewis M., Tartaglia L.A., Lee A., Bennett G.L., Rice G.C.,
RA Wong G.H., Chen E.Y., Goeddel D.V.;
RT "Cloning and expression of cDNAs for two distinct murine tumor
RT necrosis factor receptors demonstrate one receptor is species
RT specific."
RL Proc. Natl. Acad. Sci. U.S.A. 88:2830-2834(1991).
RN [2]

RP SEQUENCE FROM N.A.
RX MEDLINE-91246168; PubMed-1645445;
RA Goodwin R.G., Anderson D., Jerzy R., Davis T., Brannan C.I.,
RA Copeland N.G., Jenkins N.A., Smith C.A.;
RT "Molecular cloning and expression of the type 1 and type 2 murine
RT receptors for tumor necrosis factor."
RL Mol. Cell. Biol. 11:3020-3026(1991).
RN [3]
RP SEQUENCE OF 1-26 FROM N.A.
RC STRAIN-NOD;
RA Jacob C.O., Liu J.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 1-22 FROM N.A.
RC TISSUE=LIVER;
RA Kissoneghis M., Fellows R., Feldmann M., Chernaiovsky Y.;
RL submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: RECEPTOR FOR TNF-ALPHA.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M60469; AAA39752.1; -
DR EMBL; M59378; AAA40463.1; -
DR EMBL; U39488; AAA85021.1; -
DR EMBL; X87128; CAA60618.1; -
DR PIR; B38634; B38634.
DR HSSP; P19438; INCF.
DR MGD; MGI:1314883; TNFRSF1B.
DR INTERPRO: IPR001368; -
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00505; TNFR_NGFR_2; 3.
KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 22
FT CHAIN 23 474 TUMOR NECROSIS FACTOR RECEPTOR 2.
FT DOMAIN 23 258 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 259 288 POTENTIAL.
FT DOMAIN 289 474 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 39 203 4 X TNFR-CYS.
FT REPEAT 39 77 TNFR-CYS 1.
FT REPEAT 78 119 TNFR-CYS 2.
FT REPEAT 120 164 TNFR-CYS 3.
FT REPEAT 165 203 TNFR-CYS 4.
FT REPEAT 40 54 BY SIMILARITY.
FT DISULFID 55 68 BY SIMILARITY.
FT DISULFID 58 76 BY SIMILARITY.
FT DISULFID 79 94 BY SIMILARITY.
FT DISULFID 97 111 BY SIMILARITY.
FT DISULFID 101 119 BY SIMILARITY.
FT DISULFID 121 127 BY SIMILARITY.
FT DISULFID 136 145 BY SIMILARITY.
FT DISULFID 139 163 BY SIMILARITY.
FT DISULFID 166 181 BY SIMILARITY.
FT CARBOHYD 69 69 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 474 AA; 50319 MW; 462AE398C4D6563 CRC64;

Query Match 34.5%; Score 51; DB 1; Length 474;
Best Local Similarity 37.5%; Pred. No. 11;
Matches 9; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 1 FVQPCRRDSPTTCGPCPPPHYTQ 24
Db 63 YVKHFCNKNTSDTVCADCEASMYTQ 86

```
RESULT 6
VL1_HPV53
ID VL1_HPV53 STANDARD; PRT: 499 AA.
AC Q05113;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MAJOR CAPSID PROTEIN L1.
GN L1.
OS Human papillomavirus type 53.
OC Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Papillomavirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94265501; PubMed=8205838;
RA Delius H., Hofmann B.;
RT "Primer-directed sequencing of human papillomavirus types.";
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
[2]
RP SEQUENCE OF 300-343 FROM N.A.
RX MEDLINE=92407963; PubMed=1326639;
RA Chan S.Y., Bernard H.U., Ong C.K., Chan S.P., Birgit H., Delius H.;
RT "Phylogenetic analysis of 48 papillomavirus types and 28 subtypes and
variants: a showcase for the molecular evolution of DNA viruses.";
RL J. Virol. 66:5714-5725(1992).
[3]
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
[4]
DR EMBL; X74482; CAA52595.1; -.
DR EMBL; M96298; AAA47037.1; -.
DR PIR; S36531; S36531.
DR INTERPRO; IPR002210; -.
DR PFAM; PF00500; late_protein.L1; 1.
DR PRINTS; PR00865; HPV_CAPSID_L1.
KW Coat protein; Late protein.
SQ SEQUENCE 499 AA; 55722 MW; 5221961A3FDD5A66 CRC64;

Query Match 34.5%; Score 51; DB 1; Length 499;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 8 RDSPTTCGCP 19
DB 174 RSTPTTAGDCPP 185

RESULT 7
HAIR_MOUSE
ID HAIR_MOUSE STANDARD; PRT: 1182 AA.
AC Q61645;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE HAIRLESS PROTEIN.
GN HR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BALB/C; TISSUE=SKIN;
RX MEDLINE=94329587; PubMed=8052649;
RA Cachon-Gonzalez M.B., Fenner S., Coffin J.M., Moran C., Best S.,
RA Stoye J.P.;
RT "Structure and expression of the hairless gene of mice.";

Proc. Natl. Acad. Sci. U.S.A. 91:7717-7721(1994).
-!- FUNCTION: MAY ACT AS A TRANSCRIPTION FACTOR THAT COULD ACT ON TO
REGULATE ONE OF THE PHASES OF HAIR GROWTH.
-!- SUBCELLULAR LOCATION: NUCLEAR.
-!- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN BRAIN, HAIR
FOLLICLES AND INTERFOLLICULAR EPIDERMIS. NO EXPRESSION IN
DERMIS.
-!- DISEASE: HR MUTATION PRODUCES A NUMBER OF PLEIOTROPIC EFFECTS
INCLUDING STRUCTURAL ABNORMALITIES OF EPITHELIAL CELLS IN THE HAIR
FOLLICLES, HAIR LOSS TOWARDS THE END OF THE FIRST HAIR GROWTH
CYCLE, AND THE FAILURE OF SUBSEQUENT HAIR GROWTH CYCLES. OLDER
MICE CARRYING AN HR MUTATION HAVE BEEN REPORTED TO POSSESS ALTERED
RATIOS OF T-CELL-DEPENDENT B-CELL RESPONSES. MICE HOMOZYGOUS FOR
HR MUTATION ARE UNIQUELY SENSITIVE TO UV AND CHEMICALLY INDUCED
SKIN TUMORS.
[5]
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
[6]
DR EMBL; Z32675; CAA83587.1; -.
DR MGD; MGI:96223; HR.
KW Zinc-finger; DNA-binding; Nuclear protein; Transcription regulation;
Metal-binding.
FT ZNFING 595 620 C6-TYPE.
FT DOMAIN 535 540 POLY-GLY.
SQ SEQUENCE 1182 AA; 127182 MW; EFB80EE62AE81B40 CRC64;

Query Match 34.5%; Score 51; DB 1; Length 1182;
Best Local Similarity 34.2%; Pred. No. 25;
Matches 13; Conservative 4; Mismatches 3; Indels 18; Gaps 3;

QY 5 PCRR--DSPT--TCGP-----CPRHYTO 24
DB 338 PCRQDSPEGGSGGESSEERNKADSRACPPSHHTK 375

RESULT 8
VL1_PAPVE
ID VL1_PAPVE STANDARD; PRT: 501 AA.
AC P11326;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MAJOR CAPSID PROTEIN L1.
GN L1.
OS European elk papillomavirus (EPPV).
OC Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Papillomavirus.
RN [1]
RP SEQUENCE FROM N.A.
RA Eriksson A.;
RL Submitted (AUG-1987) to the EMBL/GenBank/DBJ databases.
[7]
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
[8]
DR EMBL; M15953; AAA66861.1; -.
DR PIR; G29499; P1WLEP.
DR INTERPRO; IPR002210; -.
DR PFAM; PF00500; late_protein.L1; 1.
DR PRINTS; PR00865; HPV_CAPSID_L1.
KW Coat protein; Late protein.
SQ SEQUENCE 501 AA; 56459 MW; 7452CCCB4964A6F6 CRC64;
```

```

Query Match          34.1%; Score 50.5; DB 1; Length 501;
Best Local Similarity 62.5%; Pred. No. 14;
Matches 10; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 4 RPCRDRSPTTCGCPPT 19
    ||| | | | | | |
Db 171 RPCRDRPET-GSCP 185

RESULT 9
VA53_VACCC STANDARD; PRT; 103 AA.
ID VA53_VACCC STANDARD; PRT; 103 AA.
AC P21071;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE PROTEIN A53.
GN A53R.
OS Vaccinia virus (strain Copenhagen).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91021027; PubMed=2219722;
RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
RA Paolletti E.;
RT "The complete DNA sequence of vaccinia virus.";
RL Virology 179:247-266(1990).
RN [2]
RP COMPLETE GENOME.
RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
RA Paolletti E.;
RL Virology 179:517-563(1990).
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M35027; AAA48188.1;
CC PIR: A42523; A42523.
CC INTERPRO: IPR001368;
CC PFAM: PF00020; TNFR_C6; 1.
CC PROSITE: PS00652; TNFR_NGFR_1; 1.
CC PROSITE: PS50050; TNFR_NGFR_2; 1.
CC Late protein; Repeat.
CC DOMAIN 36 103 2 X TNFR-CYS.
CC REPEAT 36 73 TNFR-CYS 1.
CC REPEAT 74 103 TNFR-CYS 2 (INCOMPLETE).
CC SEQUENCE 103 AA; 12032 MW; 87C3C68B5D1289CD CRC64;

Query Match          33.8%; Score 50; DB 1; Length 103;
Best Local Similarity 39.1%; Pred. No. 3.9;
Matches 9; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

QY 1 FVQRPCRRDSPTTCGCPPT 23
    | : | : | | | | |
Db 59 FAKVRCNGNDTKCERCPT 81

RESULT 10
VA53_VACCC STANDARD; PRT; 103 AA.
ID VA53_VACCC STANDARD; PRT; 103 AA.
AC P24756;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)

QY 1 FVQRPCRRDSPTTCGCPPT 23
    | : | : | | | | |
Db 59 FAKVRCNGNDTKCERCPT 81

RESULT 10
VA53_VACCC STANDARD; PRT; 103 AA.
ID VA53_VACCC STANDARD; PRT; 103 AA.
AC P24756;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)

```

```

DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE PROTEIN A53.
GN A53R OR SALF16R OR SALF19R.
OS Vaccinia virus (strain WR).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91259063; PubMed=2045793;
RA Smith G.L., Chan Y.S., Howard S.T.;
RT "Nucleotide sequence of 42 kbp of vaccinia virus strain WR from near
RT the right inverted terminal repeat.";
RL J. Gen. Virol. 72:1349-1376(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91111982; PubMed=1846491;
RA Howard S.T., Chan Y.S., Smith G.L.;
RT "Vaccinia virus homologues of the Shope fibroma virus inverted
RT terminal repeat proteins and a discontinuous ORF related to the tumor
RT necrosis factor receptor family.";
RL Virology 180:633-647(1991).
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D11079; BAA01827.1;
CC PIR: B38550; AAA48339.1;
CC PIR: B38550; B38550.
CC PIR: JQ1791; JQ1791.
CC INTERPRO: IPR001368;
CC PFAM: PF00020; TNFR_C6; 1.
CC PROSITE: PS00652; TNFR_NGFR_1; 1.
CC PROSITE: PS50050; TNFR_NGFR_2; 1.
CC Late protein; Repeat.
CC DOMAIN 36 103 2 X TNFR-CYS.
CC REPEAT 36 73 TNFR-CYS 1.
CC REPEAT 74 103 TNFR-CYS 2 (INCOMPLETE).
CC SEQUENCE 103 AA; 12001 MW; 87C3C68B5D1339CD CRC64;

Query Match          33.8%; Score 50; DB 1; Length 103;
Best Local Similarity 39.1%; Pred. No. 3.9;
Matches 9; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

QY 1 FVQRPCRRDSPTTCGCPPT 23
    | : | : | | | | |
Db 59 FAKVRCNGNDTKCERCPT 81

RESULT 11
US02_HSVEB STANDARD; PRT; 418 AA.
ID US02_HSVEB STANDARD; PRT; 418 AA.
AC P28964;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE GENE 68 PROTEIN.
GN 68.
OS Equine herpesvirus type 1 (strain Ab4p) (EHV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92295566; PubMed=1318606;
RA Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
RT "The DNA sequence of equine herpesvirus-1.";
RL Virology 189:304-316(1992).

```


GN CFH.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86233353; PubMed=2940596;
 RA Kristensen T., Tack B.F.;
 RT "Murine protein H is comprised of 20 repeating units, 61 amino acids
 in length.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:3963-3967(1986).
 RN [2]
 RP SEQUENCE OF 1-19 FROM N.A.
 RC STRAIN=BALB/C;
 RX MEDLINE=90148935; PubMed=2533512;
 RA Munoz-Canoves P., Tack B.F., Vik D.P.;
 RT "Analysis of complement factor H mRNA expression: dexamethasone and
 IFN-gamma increase the level of H in L cells.";
 RL Biochemistry 28:9891-9897(1989).
 RN [3]
 RP SEQUENCE OF 1-18 FROM N.A.
 RX MEDLINE=90111033; PubMed=2136885;
 RA Natsume-Sakai S., Nonaka M., Nonaka Y.N., Shreffler D.C.,
 RA Moriwaki K.;
 RT "Demonstration of an unusual allelic variation of mouse factor H by
 the complete cDNA sequence of the H.2 allotype.";
 RL J. Immunol. 144:358-362(1990).
 CC -1- FUNCTION: FACTOR H FUNCTIONS AS A COFACTOR IN THE INACTIVATION OF
 C3B BY FACTOR I AND ALSO INCREASES THE RATE OF DISSOCIATION OF THE
 C3BBB COMPLEX (C3 CONVERTASE) AND THE (C3b)NBB COMPLEX (C5
 CONVERTASE) IN THE ALTERNATIVE COMPLEMENT PATHWAY.
 CC -1- POLYMORPHISM: TWO CODOMINANT ALLELES OF FACTOR H ARE PRESENT IN
 MICE.
 CC -1- SIMILARITY: CONTAINS 20 SUSHI (SCR) REPEATS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M12660; AAA37759.1; -;
 DR EMBL; J02891; AAA37795.1; -;
 DR EMBL; M31979; AAA37762.1; -;
 DR PIR; A26154; NRMSH.
 DR HSP; P08603; IHFI.
 DR MGD; MGI:88385; CFH.
 DR INTERPRO; IPR000436; -;
 DR PFAM; PF00084; sushi; 20.
 KW Complement alternate pathway; Plasma; Glycoprotein; Repeat; Sushi;
 Signal.
 FT SIGNAL
 FT CHAIN 1 18 BY SIMILARITY
 FT DOMAIN 19 1234 COMPLEMENT FACTOR H
 FT REPEAT 20 1224 20 x SUSHI (SCR) REPEATS.
 FT REPEAT 20 81 SUSHI 1.
 FT REPEAT 84 142 SUSHI 2.
 FT REPEAT 145 206 SUSHI 3.
 FT REPEAT 209 263 SUSHI 4.
 FT REPEAT 266 321 SUSHI 5.
 FT REPEAT 324 386 SUSHI 6.
 FT REPEAT 388 443 SUSHI 7.
 FT REPEAT 447 506 SUSHI 8.
 FT REPEAT 508 565 SUSHI 9.
 FT REPEAT 568 623 SUSHI 10.
 FT REPEAT 628 684 SUSHI 11.
 FT REPEAT 689 744 SUSHI 12.
 FT REPEAT 751 803 SUSHI 13.
 FT REPEAT 807 862 SUSHI 14.
 FT REPEAT 866 932 SUSHI 15.
 FT REPEAT 935 990 SUSHI 16.
 FT REPEAT 993 1049 SUSHI 17.

FT REPEAT 1052 1108
 FT REPEAT 1113 1169
 FT REPEAT 1171 1234
 FT DISULFID 21 66 BY SIMILARITY.
 FT DISULFID 52 80 BY SIMILARITY.
 FT DISULFID 85 129 BY SIMILARITY.
 FT DISULFID 114 141 BY SIMILARITY.
 FT DISULFID 146 192 BY SIMILARITY.
 FT DISULFID 178 205 BY SIMILARITY.
 FT DISULFID 210 251 BY SIMILARITY.
 FT DISULFID 237 262 BY SIMILARITY.
 FT DISULFID 267 309 BY SIMILARITY.
 FT DISULFID 294 320 BY SIMILARITY.
 FT DISULFID 325 374 BY SIMILARITY.
 FT DISULFID 357 385 BY SIMILARITY.
 FT DISULFID 389 431 BY SIMILARITY.
 FT DISULFID 416 442 BY SIMILARITY.
 FT DISULFID 448 494 BY SIMILARITY.
 FT DISULFID 477 505 BY SIMILARITY.
 FT DISULFID 509 553 BY SIMILARITY.
 FT DISULFID 536 564 BY SIMILARITY.
 FT DISULFID 569 610 BY SIMILARITY.
 FT DISULFID 597 622 BY SIMILARITY.
 FT DISULFID 629 672 BY SIMILARITY.
 FT DISULFID 658 683 BY SIMILARITY.
 FT DISULFID 690 732 BY SIMILARITY.
 FT DISULFID 718 743 BY SIMILARITY.
 FT DISULFID 752 791 BY SIMILARITY.
 FT DISULFID 780 802 BY SIMILARITY.
 FT DISULFID 808 850 BY SIMILARITY.
 FT DISULFID 836 861 BY SIMILARITY.
 FT DISULFID 867 920 BY SIMILARITY.
 FT DISULFID 906 931 BY SIMILARITY.
 FT DISULFID 936 978 BY SIMILARITY.
 FT DISULFID 964 989 BY SIMILARITY.
 FT DISULFID 994 1037 BY SIMILARITY.
 FT DISULFID 1023 1048 BY SIMILARITY.
 FT DISULFID 1053 1096 BY SIMILARITY.
 FT DISULFID 1082 1107 BY SIMILARITY.
 FT DISULFID 1114 1157 BY SIMILARITY.
 FT DISULFID 1143 1168 BY SIMILARITY.
 FT DISULFID 1172 1223 BY SIMILARITY.
 FT DISULFID 1206 1233 BY SIMILARITY.
 FT CARBOHYD 676 676 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 721 721 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 773 773 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 801 801 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1030 1030 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1061 1061 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1225 1225 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1234 AA; 139082 MW; C5AC02F341B957F7 CRC64;

Query Match 33.8%; Score 50; DB 1; Length 1234;
 Best Local Similarity 47.4%; Pred. No. 35;
 Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 FVQRPCRRDSPTTCGRCPP 19
 : : : : :
 Db 1101 WTEKPCRDSTGKCGPPP 1119

RESULT 14
 EFS_HUMAN
 ID EFS_HUMAN STANDARD; PRT; 561 AA.
 AC O43281; O43282;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE EMBRYONAL FYN-ASSOCIATED SUBSTRATE (HEFS).
 GN EFS.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HIPPOCAMPUS;
RX MEDLINE=98007065; PubMed=9349509;
RA Ishino M., Ohba T., Inazawa J., Sasaki H., Ariyama Y., Sasaki T.;
RT "Identification of an Efs isoform that lacks the SH3 domain and
RT chromosomal mapping of human Efs.";
RL Oncogene 15:1741-1745(1997).
CC -!- FUNCTION: DOCKING PROTEIN WHICH PLAYS A CENTRAL COORDINATING ROLE
CC FOR TYROSINE-KINASE-BASED SIGNALING RELATED TO CELL ADHESION. MAY
CC SERVE AS AN ACTIVATOR OF SRC AND A DOWNSTREAM EFFECTOR. INTERACTS
CC WITH THE SH3 DOMAIN OF FYN AND WITH CRK, SRC, AND YES (BY
CC SIMILARITY).
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS: ISOFORM EFS1 (SHOWN HERE)
CC ISOFORM EFS2; MAY BE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: THE PROTEIN HAS BEEN DETECTED IN LUNG AND
CC PLACENTA.
CC -!- DOMAIN: CONTAINS A CENTRAL DOMAIN (SUBSTRATE DOMAIN) CONTAINING
CC MULTIPLE POTENTIAL SH2-BINDING SITES AND A C-TERMINAL DOMAIN
CC CONTAINING A DIVERGENT HELIX-LOOP-HELIX (HLH) MOTIF. THE SH2-
CC BINDING SITES PUTATIVELY BIND CRK, NCK AND ABL SH2 DOMAINS.
CC -!- PTM: PHOSPHORYLATED ON MULTIPLE TYROSINE RESIDUES. PHOSPHORYLATED
CC ON TYROSINES BY FYN AND SRC (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN. THE SH3 DOMAIN INTERACTS WITH
CC FOCAL ADHESION KINASE 1.
CC -!- SIMILARITY: CONTAINS 2 SH3-BINDING DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE CAS FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; AB001466; BAA24588.1; -
DR EMBL; AB001467; BAA24589.1; -
DR HSP; P29354; IGRI.
DR INTERPRO: IPR001452; -
DR PFAM: PF00018; SH3; 1.
DR PROSITE: PS50002; SH3; 1.
KW Phosphorylation; SH3 domain; SH3-binding; Cell adhesion;
KW Alternative splicing.
FT DOMAIN 5 68
FT DOMAIN 104 124
FT DOMAIN 288 291
FT DOMAIN 305 311
FT DOMAIN 335 341
FT DOMAIN 336 342
FT DOMAIN 438 488
FT MOD_RES 253 253
FT VARSPPLIC 7 99
FT SEQUENCE 561 AA; 58815 MW; 30FFFIAD4C9D1C4A CRC64;
Query Match 33.4%; Score 49.5; DB 1; Length 561;
Best Local Similarity 56.2%; Pred. No. 20;
Matches 9; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
QY 4 RPCRDSPTTCGPP 19
DB 108 RCPPTSGP-PAGPCPP 122
RESULT 15
OX40_HUMAN STANDARD; PRT; 277 AA.
AC P43489; Q13663;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE OX40L RECEPTOR PRECURSOR (ACT35 ANTIGEN) (TAX-TRANSCRIPTIONALLY
GN ACTIVATED GLYCOPROTEIN 1 RECEPTOR) (CD134 ANTIGEN).
GN TNFRSF4 OR TAGP1L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94170844; PubMed=7510240;
RA Latza U., Duerkop H., Schnittger S., Ringeling J., Eitelbach F.,
RA Hummel M., Fonatsch C., Stein H.;
RT "The human OX40 homolog: cDNA structure, expression and chromosomal
RT assignment of the ACT35 antigen.";
RL Eur. J. Immunol. 24:677-683(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95219871; PubMed=7704935;
RA Baum P.R., Gayle R.B. III, Ramsdell F., Srinivasan S., Sorensen R.A.,
RA Watson M.L., Seldin M.F., Clifford K.N., Grabstein K., Alderson M.R.;
RT "Identification of OX40 ligand and preliminary characterization of
RT its activities on OX40 receptor.";
RL Circ. Shock 44:30-34(1994).
CC -!- FUNCTION: RECEPTOR FOR THE OX40L/GP34 CYTOKINE.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD134 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd134.htm".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; X75962; CAA53576.1; -
DR EMBL; S76792; AAB33944.1; ALT_INIT.
DR HSP; P25942; ICDF.
DR MIM; 600315; -
DR INTERPRO: IPR001368; -
DR PFAM: PF00020; TNFR_C6; 3.
DR PROSITE: PS00652; TNFR_NGFR_1; 3.
DR PROSITE: PS50050; TNFR_NGFR_2; 2.
KW Receptor; T-cell; Antigen; Glycoprotein; Transmembrane; Repeat;
KW Signal.
FT SIGNAL 1 28
FT CHAIN 29 277
FT DOMAIN 29 214
FT DOMAIN 215 235
FT TRANSMEM 236 277
FT DOMAIN 30 169
FT DOMAIN 30 65
FT REPEAT 36 107
FT REPEAT 108 126
FT REPEAT 127 167
FT REPEAT 146 146
FT CARBOHYD 146 160
FT CARBOHYD 160 160
FT SEQUENCE 277 AA; 29340 MW; 49F1525941550BF CRC64;
Query Match 33.1%; Score 49; DB 1; Length 277;
Best Local Similarity 60.0%; Pred. No. 12;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 14 CGPCPPRHYT 23
DB 125 CAPCPGHS 134
Search completed: January 30, 2001, 17:06:09
Job time: 1230 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:54:58 ; Search time 254.27 seconds
(without alignments)
11.063 Million cell updates/sec

Title: US-09-518-931-4_COPY_57_80
Perfect score: 148
Sequence: 1 FVQPCRRDPTTCGCPPRHYTQ 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues
Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_15:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	148	100.0	300	4	Q95407
2	56.5	38.2	228	4	Q9KJN5
3	56	37.8	249	5	Q9W0I5
4	56	37.8	372	4	Q90HP4
5	56	37.8	401	4	Q00300
6	55	37.2	289	5	Q62432
7	55	37.2	290	5	Q44173
8	54	36.5	709	2	Q49145
9	53.5	36.1	813	4	Q75229
10	53	35.8	299	5	Q20087
11	53	35.8	299	5	Q18302
12	53	35.8	299	5	Q27318
13	53	35.8	423	4	Q9NSV2
14	53	35.8	423	4	Q9NS68
15	52	35.1	267	6	Q02764
16	52	35.1	606	10	Q04504
17	51.5	34.8	177	10	Q9LIU5
18	51	34.5	283	4	Q92956
19	51	34.5	283	4	Q9UM65

20	51	34.5	299	5	Q20805
21	51	34.5	299	5	Q22393
22	51	34.5	459	11	Q62327
23	51	34.5	676	5	Q9VQS0
24	51	34.5	1477	4	Q95722
25	51	34.5	3014	4	Q9NYQ6
26	50.5	34.1	102	12	Q86542
27	50.5	34.1	369	5	Q9NEK8
28	50.5	34.1	438	10	Q9XEP7
29	50	33.8	103	4	Q16221
30	50	33.8	186	12	Q9WJB4
31	50	33.8	304	5	Q20739
32	50	33.8	305	5	Q17805
33	50	33.8	316	5	Q19111
34	50	33.8	327	5	Q18975
35	50	33.8	334	5	Q24403
36	50	33.8	334	5	Q9VAV8
37	50	33.8	503	5	Q9W0I4
38	50	33.8	564	10	Q9SM88
39	50	33.8	1146	4	Q95856
40	50	33.8	1203	4	Q95341
41	50	33.8	1203	4	Q9Y6C5
42	50	33.8	1764	11	Q35806
43	50	33.8	1833	11	Q08999
44	50	33.8	1963	6	Q28019
45	49.5	33.4	186	5	Q9XV17

ALIGNMENTS

RESULT 1	
Q95407	
ID	Q95407 PRELIMINARY; PRT; 300 AA.
AC	Q95407;
DT	01-MAY-1999 (TREMBLrel. 10, Created)
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE	DECOY RECEPTOR 3 (M68) (M68C) (M68E).
GN	DCR3 OR TR6.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=99087326; PubMed=9872321;
RA	Pitti R.M., Marsters S.A., Lawrence D.A., Roy M., Kischkel F.C.,
RA	Dowd P., Huang A., Donahue C.J., Sherwood S.W., Baldwin D.T.,
RA	Godowski P.J., Wood W.I., Gurney A.L., Hillan K.J., Cohen R.L.,
RA	Goddard A.D., Botstein D., Ashkenazi A.;
RT	"Genomic amplification of a decoy receptor for Fas ligand in lung and
RT	colon cancer.";
RL	Nature 396:699-703(1998).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	TISSUE=BLOOD;
RX	MEDLINE=99253915; PubMed=10318773;
RA	Yu K.Y., Kwon B., Ni J., Zhai Y., Ebner R., Kwon B.S.;
RT	"A newly identified member of tumor necrosis factor receptor
RT	superfamily (TR6) suppresses LIGHT-mediated apoptosis.";
RL	J. Biol. Chem. 274:13733-13736(1999).
RN	[3]
RP	SEQUENCE FROM N.A.
RC	TISSUE=PANCREAS;
RX	MEDLINE=20122600; PubMed=10655513;
RA	Bai C., Connolly B., Metzker M.L., Hilliard C.A., Liu X., Sandig V.,
RA	Soderman A., Galloway S.M., Liu Q., Austin C.P., Caskey C.T.;
RT	"Overexpression of M68/DCR3 in human gastrointestinal tract tumors
RT	independent of gene amplification and its location in a four-gene
RT	cluster.";
RL	Proc. Natl. Acad. Sci. U.S.A. 97:1230-1235(2000).
DR	EMBL; AF104419; AAD03056.1; -.

```

DR EMBL; AF134240; AAD29688.1; -
DR EMBL; AF217796; AAF35244.1; -
DR EMBL; AF217793; AAF33685.1; -
DR EMBL; AF217794; AAF33686.1; -
DR HSSP; P25942; ICDF;
DR INTERPRO; IPR000561; -
DR INTERPRO; IPR001368; -
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS00030; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -; 1.
KW Receptor.
SQ SEQUENCE 300 AA; 32679 MW; F90AEE33718449AF CRC64;

Query Match 100.0%; Score 148; DB 4; Length 300;
Best Local Similarity 100.0%; Pred. No. 5.4e-13;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FVQPCRRDSTTCGCPPRHYTQ 24
|||||
DB 57 FVQPCRRDSTTCGCPPRHYTQ 80

RESULT 2
Q9KJN5 PRELIMINARY; PRT; 228 AA.
AC Q9KJN5
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE PUTATIVE RESPONSE REGULATOR.
OS Myxococcus xanthus.
OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
OC Myxococcales; Cystobacterineae; Myxococcaceae; Myxococcus.
OX NCBI_TaxID=34;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D22;
RX MEDLINE=20032045; PubMed=10564511;
RA Cho K., Zusman D.R.;
RT "Sporulation timing in Myxococcus xanthus is controlled by the espB locus.";
RL Mol. Microbiol. 34:714-725(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=D22;
RA Cho K., Zusman D.R.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF163841; AAF87932.1; -
SQ SEQUENCE 228 AA; 24921 MW; B4BAD784F1E8F65C CRC64;

Query Match 38.2%; Score 56.5; DB 2; Length 228;
Best Local Similarity 52.4%; Pred. No. 1.4;
Matches 11; Conservative 1; Mismatches 8; Indels 1; Gaps 1;

QY 4 RPCRDRDSTTCGCPPRHYTQ 24
|||
DB 51 RPEARGP-RCGACPPHHKEQ 70

RESULT 3
Q9W0I5 PRELIMINARY; PRT; 249 AA.
ID Q9W0I5
AC Q9W0I5
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CG129 PROTEIN.
GN CG129.
OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Calle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballwe R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotzler P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foslere C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Smith H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Sten T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-E., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003471; AAF47461.1; -
DR FLYBASE; FBgn0035196; CG9129.
SQ SEQUENCE 249 AA; 28593 MW; 12E80A27D43CAA85 CRC64;

Query Match 37.8%; Score 56; DB 5; Length 249;
Best Local Similarity 55.6%; Pred. No. 1.8;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 VQPCRRDSTTCGCPGPP 19
|||||
DB 208 VQPCVCKDPCNCCCPGPP 225

RESULT 4
Q9UHP4 PRELIMINARY; PRT; 372 AA.
ID Q9UHP4
AC Q9UHP4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE OSTEOPTOTERIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

```

```
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA He Z.-Y., Yang G.-Z., Zhang W.-J., Wu X.-F.;
RT "Cloning and Expression of Osteoprotegerin from Homo sapiens.";
RL Sheng Wu Hua Hsueh Yu Sheng Wu Li Hsueh Pao 31:680-684(1999).
DR EMBL; AF134187; AAF20168.1; -.
DR HSSP; P25942; ICDF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
FT NON_TER 1
SQ SEQUENCE 372 AA; 42758 MW; F02527A5CD01CCD3 CRC64;

Query Match 37.8%; Score 56; DB 4; Length 372;
Best Local Similarity 34.8%; Pred. No. 2.4;
Matches 8; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 FVQPCRRDSTTCGCPPRHYT 23
Db 28 YLKQCHTAKWTVCAPCPDHYT 50
::: | | | | |
RESULT 5
000300 PRELIMINARY; PRT; 401 AA.
ID 000300; O60236;
AC 000300; O60236;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)
DE (OCIF) (TUMOR NECROSIS FACTOR RECEPTOR SUPERFAMILY MEMBER 11B).
GN TNFRSF11B OR OPG OR OCIF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RX MEDLINE=97262071; PubMed=9108485;
RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,
RA Luehly R., Nguyen H.O., Woodson S., Bennett L., Boone T., Shimamoto G.,
RA Derose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,
RA Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,
RA Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,
RA Suggs S., Boyle W.J.;
RT "Osteoprotegerin: a novel secreted protein involved in the regulation
RT of bone density.";
RL Cell 89:309-319(1997).
[2]
RN SEQUENCE FROM N.A.
RC TISSUE=LUNG FIBROBLAST;
RX MEDLINE=98151033; PubMed=9492085;
RA Yasuda H., Shima N., Nakagawa N., Mochizuki S.-I., Yano K., Fujise N.,
RA Sato Y., Goto M., Yamaguchi K., Kuriyama M., Kanno T., Murakami A.,
RA Tsuda E., Morinaga T., Higashio K.;
RT "Identity of osteoclastogenesis inhibitory factor (OCIF) and
RT osteoprotegerin (OPG): a mechanism by which OPG/OCIF inhibits
RT osteoclastogenesis in vitro.";
RL Endocrinology 139:1329-1337(1998).
[3]
RN SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=98351569; PubMed=9688283;
RA Morinaga T., Nakagawa N., Yasuda H., Tsuda E., Higashio K.;
RT "Cloning and characterization of the gene encoding human
RT osteoprotegerin/osteoclastogenesis-inhibitory factor.";
RL Eur. J. Biochem. 254:685-691(1998).
CC -1- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES
CC OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY
CC SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO
```

```
CC OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN
CC STROMAL CELLS AND OSTEOCLAST PROGENITORS.
CC -1- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM) (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGHEST LEVELS IN LUNG, HEART,
CC KIDNEY, PLACENTA, THYROID, SPINAL CORD AND LIVER. ALSO DETECTED IN
CC A NUMBER OF OTHER HEMATOPOIETIC AND IMMUNE ORGANS. NOT DETECTED IN
CC THE PANCREAS OR PERIPHERAL BLOOD LYMPHOCYTES.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL; AB002146; BAA35910.1; -.
DR EMBL; AB008822; BAA32076.1; -.
DR EMBL; AB008821; BAA32076.1; JOINED.
DR EMBL; U94332; AAB53709.1; -.
DR HSSP; P25942; ICDF.
DR MIM; 602643; -.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -.
KW Glycoprotein; Repeat; Cytokine; Signal.
FT SIGNAL 1 21 BY SIMILARITY.
FT CHAIN 22 401 OSTEOPROTEGERIN.
FT DOMAIN 23 183 4 X TNFR-CYS.
FT REPEAT 23 63 TNFR-CYS 1.
FT REPEAT 64 106 TNFR-CYS 2.
FT REPEAT 107 143 TNFR-CYS 3.
FT REPEAT 144 201 TNFR-CYS 4.
FT DOMAIN 306 365 DEATH DOMAIN.
FT DISULFID 41 54 BY SIMILARITY.
FT DISULFID 44 62 BY SIMILARITY.
FT DISULFID 65 80 BY SIMILARITY.
FT DISULFID 83 97 BY SIMILARITY.
FT DISULFID 87 105 BY SIMILARITY.
FT DISULFID 118 142 BY SIMILARITY.
FT DISULFID 145 160 BY SIMILARITY.
FT CARBOHYD 98 98 POTENTIAL.
FT CARBOHYD 152 152 POTENTIAL.
FT CARBOHYD 165 165 POTENTIAL.
FT CARBOHYD 178 178 POTENTIAL.
FT CARBOHYD 289 289 POTENTIAL.
FT CONFLICT 263 263 A -> D (IN REF. 2 AND 3).
SQ SEQUENCE 401 AA; 45996 MW; EB42FA51C9D7C71E CRC64;

Query Match 37.8%; Score 56; DB 4; Length 401;
Best Local Similarity 34.8%; Pred. No. 2.6;
Matches 8; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 1 FVQPCRRDSTTCGCPPRHYT 23
Db 49 YLKQCHTAKWTVCAPCPDHYT 71
::: | | | | |
RESULT 6
062432 PRELIMINARY; PRT; 289 AA.
ID 062432
AC 062432;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE Y41E3.2 PROTEIN.
GN Y41E3.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurray A.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
```

RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA Snaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans.";
RL Nature 368:32-38(1994).
DR EMBL: Z95559; CAB08999.1; -.
DR INTERPRO: IPR000087; -.
DR INTERPRO: IPR002486; -.
DR PFAM: PF01391; Collagen; 2.
DR PFAM: PF01484; Col_cuticle_N; 1.
SQ SEQUENCE 289 AA; 28777 MW; CB4E0ADOC1551364 CRC64;

Query Match 37.2%; Score 55; DB 5; Length 289;
Best Local Similarity 50.0%; Pred. No. 2.7;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QRCRRDSPPTCGPCP 18
|:|:| :| :| :| :|
Db 124 QQCEPTTPPCQPCP 139

RESULT 7
Q44173 PRELIMINARY; PRT; 290 AA.
AC Q44173
DT 01-JUN-1998 (TEMBLrel. 06, Created)
DT 01-JUN-1998 (TEMBLrel. 06, Last sequence update)
DT 01-OCT-2000 (TEMBLrel. 15, Last annotation update)
DE F58F6.2 PROTEIN.
GN F58F6.2
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA Snaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans.";
RL Nature 368:32-38(1994).
DR EMBL: Z95559; CAB08999.1; -.
DR INTERPRO: IPR000087; -.
DR INTERPRO: IPR002486; -.
DR PFAM: PF01391; Collagen; 2.
DR PFAM: PF01484; Col_cuticle_N; 1.
SQ SEQUENCE 289 AA; 28777 MW; CB4E0ADOC1551364 CRC64;

DR INTERPRO: IPR000087; -.
DR INTERPRO: IPR002486; -.
DR PFAM: PF01391; Collagen; 2.
DR PFAM: PF01484; Col_cuticle_N; 1.
SQ SEQUENCE 290 AA; 29279 MW; 9CA8218599A18A87 CRC64;

Query Match 37.2%; Score 55; DB 5; Length 290;
Best Local Similarity 50.0%; Pred. No. 2.8;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QRCRRDSPPTCGPCP 18
|:|:| :| :| :| :|
Db 125 QQCELVTPPPCPCP 140

RESULT 8
Q49145 PRELIMINARY; PRT; 709 AA.
AC Q49145
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TEMBLrel. 13, Last annotation update)
DE PROTEASE.
GN PQQE.
OS Methylobacterium extorquens.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Methylobacterium group; Methylobacterium.
OX NCBI_TaxID=408;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AM1;
RX MEDLINE=96178988; PubMed=8606199;
RA Springer A.L., Ramamoorthi R., Lidstrom M.E.;
RT "Characterization and nucleotide sequence of pqqE and pqqF in
Methylobacterium extorquens AM1.";
RL J. Bacteriol. 178:2134-2157(1996).
DR EMBL: L43135; AAB00962.1; -.
DR INTERPRO: IPR001431; -.
DR PFAM: PF00675; Peptidase M16; 1.
DR PROSITE: PS00143; INSULINASE; 1.
KW Protease.
SQ SEQUENCE 709 AA; 77127 MW; 41659464578F05B6 CRC64;

Query Match 36.5%; Score 54; DB 2; Length 709;
Best Local Similarity 64.3%; Pred. No. 7.5;
Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 RPCRDRSPPTCGPCP 17
|:|:| :| :| :| :|
Db 693 RPCRDRSPPTSSPC 706

RESULT 9
O75229 PRELIMINARY; PRT; 813 AA.
AC O75229
DT 01-NOV-1998 (TEMBLrel. 08, Created)
DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)
DT 01-MAY-2000 (TEMBLrel. 13, Last annotation update)
DE R31449_3 (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E., McCready P.M., Skowronski E., Adamson A.W.,
RA Burkhardt-Schultz K., Gordon L., Kyle A., Ramirez M., Stilwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Barnes J.,
RA Danganan L., Poundstone P., Christensen M., Georgescu A., Avila J.,
RA Liu S., Attix C., Andreise T., Trankheim M., Amico-Keller G.,

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE T28C6.6 PROTEIN.
GN T28C6.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA LLOYD C.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favellio A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopka A., Saunders D., Shownkeen R.,
RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstein L., Wilkinson-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
DR ENBL: 254238; CAA90997.1; -;
DR ENBL: 254238; CAA90995.1; -;
DR INTERPRO: IPR000087; -;
DR PFAM: PF01391; Collagen; 2.
DR PFAM: PF01484; Col.cuticle.N; 1.
SQ SEQUENCE 299 AA; 28660 MW; 50C03F1E1D26A6D0 CRC64;

Query Match 35.8%; Score 53; DB 5; Length 299;
Best Local Similarity 50.0%; Pred. No. 5.3;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QRCRRDSPTTCGCP 18
|:|:| :| :| :|
Db 134 QQCPEPTPPPCPKPC 149

RESULT 13

ID Q9NZV2 PRELIMINARY; PRT; 423 AA.
AC Q9NZV2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE TAJ-ALPHA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Eby M.T., Jasmin A., Kumar A., Sharma K., Chaudhary P.M.;
RT "TAJ, a Novel Member of the Tumor Necrosis Factor Receptor Family,
RT Activates the c-Jun N-terminal Kinase Pathway and Mediates Caspase-
RT Independent Cell Death.";
RL J. Biol. Chem. 275:15336-15342(2000).
DR ENBL: AFI67555; AAF71828.1; -;
SQ SEQUENCE 423 AA; 46071 MW; BB7C9917132A4B2F CRC64;

Query Match 35.8%; Score 53; DB 4; Length 423;
Best Local Similarity 81.8%; Pred. No. 6.9;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 CRRDSTTCGP 16
|:|:| :| :| :|
Db 237 CRRDSVQTCGP 247

RESULT 14

ID Q9NS68 PRELIMINARY; PRT; 423 AA.
AC Q9NS68;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE HTROY.
GN HTROY.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kojima T., Morikawa Y., Copeland N.G., Gilbert D.J., Jenkins N.A.,
RA Senba E., Kitamura T.;
RT "TROY, a newly identified member of the tumor necrosis factor receptor
RT superfamily, exhibits a homology with Edar and is expressed in
RT embryonic skin and hair follicles.";
RL J. Biol. Chem. 275:20742-20747(2000).
DR ENBL: AB040434; BAB03269.1; -;
SQ SEQUENCE 423 AA; 46015 MW; B5891CEA9ED45361 CRC64;

Query Match 35.8%; Score 53; DB 4; Length 423;
Best Local Similarity 81.8%; Pred. No. 6.9;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 CRRDSTTCGP 16
|:|:| :| :| :|
Db 237 CRRDSVQTCGP 247

RESULT 15

ID O02764 PRELIMINARY; PRT; 267 AA.
AC O02764;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE OX40 PRECURSOR (FRAGMENT).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CHBB:HM;
RA Isono T., Seto A.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR ENBL: AB003911; BAA20059.1; -;
DR HSSP: P19438; 1EXT.
DR INTERPRO: IPR001368; -;
DR PFAM: PF000020; TNFR_c6; 3.
DR PROSITE: PS00652; TNFR_NGFR_1; 2.
DR PROSITE: PS50050; TNFR_NGFR_2; 2.
KW Signal.
FT NON_TER 1 1
FT SIGNAL <1 18 POTENTIAL.
FT CHAIN 19 267 OX40.
SQ SEQUENCE 267 AA; 28489 MW; A8B4CD3173C9500B CRC64;

Query Match 35.1%; Score 52; DB 6; Length 267;
Best Local Similarity 42.1%; Pred. No. 6.6;
Matches 8; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

Qy 6 CRRDPTTCGCPPRHYTQ 24
| | | | | | | | :
Db 50 CNRSODTICHPCPGFYNE 68

Search completed: January 30, 2001, 16:55:00
Job time: 916 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:47:48 ; Search time 132.77 seconds
(without alignments)
3.246 Million cell updates/sec

Title: US-09-518-931-4_COPY_57_80

Perfect score: 148

Sequence: 1 FVQPCRRDSTTCGCPPRHYTQ 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_AA:*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
- 3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
- 4: /cgn2_6/ptodata/2/iaa/PCITUS_COMB.pep:*
- 5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	148	100.0	300	2	US-08-794-796-2
2	56	37.8	401	3	US-08-974-022-6
3	53	35.8	38	1	US-08-239-256-16
4	51	34.5	283	4	US-08-12374-2
5	51	34.5	474	2	US-08-650-000-4
6	51	34.5	474	5	US-08-650-000-4
7	50	33.8	416	1	US-08-117-083-62
8	50	33.8	967	2	US-08-449-645A-30
9	50	33.8	967	2	US-08-702-367A-30
10	50	33.8	991	2	US-08-449-645A-13
11	50	33.8	991	2	US-08-702-367A-13
12	50	33.8	991	4	US-08-449-645A-13
13	50	33.8	1833	3	US-08-479-722B-2
14	50	33.8	1833	4	US-08-479-722B-2
15	49	33.1	277	2	US-08-147-784-2
16	49	33.1	755	3	US-09-071-101-2
17	49	33.1	755	3	US-09-369-618-2
18	49	33.1	755	3	US-09-369-617-2
19	48.5	32.8	627	2	US-08-466-589-6
20	48.5	32.8	627	2	US-08-700-636-6
21	48.5	32.8	627	3	US-08-467-574-6
22	48	32.4	45	1	US-08-117-083-19
23	48	32.4	139	2	US-08-219-237B-8
24	48	32.4	186	1	US-08-089-458B-6
25	48	32.4	205	3	US-08-974-022-51
26	48	32.4	401	3	US-08-974-022-4
27	48	32.4	928	1	US-08-442-248-2
28	48	32.4	928	1	US-08-440-815-2

29	48	32.4	1005	2	US-08-469-537A-103	Sequence 103, Appl
30	47	31.8	206	1	US-08-097-827-7	Sequence 7, Appli
31	47	31.8	206	1	US-08-494-574-7	Sequence 7, Appli
32	47	31.8	438	1	US-08-097-827-11	Sequence 11, Appl
33	47	31.8	438	1	US-08-494-574-11	Sequence 11, Appl
34	46.5	31.4	396	1	US-07-649-591B-4	Sequence 4, Appli
35	46.5	31.4	396	1	US-08-277-540-4	Sequence 4, Appli
36	46.5	31.4	396	1	US-08-430-787A-4	Sequence 4, Appli
37	46.5	31.4	1104	2	US-08-327-832-5	Sequence 5, Appli
38	46.5	31.4	1104	2	US-08-828-584-5	Sequence 5, Appli
39	46.5	31.4	1251	4	PCT-US95-02251-3	Sequence 3, Appli
40	46.5	31.4	1252	1	US-08-199-780-3	Sequence 3, Appli
41	46.5	31.4	1252	2	US-08-316-650-3	Sequence 3, Appli
42	46.5	31.4	1253	3	US-08-479-722B-4	Sequence 4, Appli
43	46.5	31.4	2050	2	US-08-347-594A-2	Sequence 2, Appli
44	46	31.1	48	3	US-08-974-022-43	Sequence 43, Appl
45	46	31.1	70	3	US-08-974-022-41	Sequence 41, Appl

ALIGNMENTS

RESULT 1
US-08-794-796-2
; Sequence 2, Application US/08794796
; Patent No. 5885800
; GENERAL INFORMATION:
; APPLICANT: Emery, John
; APPLICANT: Tan, KB
; APPLICANT: Truneh, Alem
; APPLICANT: Young, Peter
; TITLE OF INVENTION: Tumor Necrosis Related Receptor,
; TITLE OF INVENTION: TR4
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESS: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/794,796
; FILING DATE: 04-FEB-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: GH50000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-4026
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 300 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-794-796-2

Query Match 100.0%; Score 148; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 5.8e-10;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FVORPCRRDSPTTCGCPPPRHYYTQ 24
|||||
Db 57 FVORPCRRDSPTTCGCPPPRHYYTQ 80

RESULT 2

US-08-974-022-6
; Sequence 6, Application US/08974022
; Patent No. 6015938
; GENERAL INFORMATION:
; APPLICANT: Boyle, William J.
; APPLICANT: Lacey, David L.
; APPLICANT: Calzone, Frank J.
; APPLICANT: Chang, Ming-Shi
; TITLE OF INVENTION: OSTEOPROTEGERIN
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 12-DEC-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/577,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-378
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 401 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-974-022-6

Query Match 37.8%; Score 56; DB 3; Length 401;
Best Local Similarity 34.8%; Pred. No. 16;
Matches 8; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Qy 1 FVORPCRRDSPTTCGCPPPRHYYT 23
:::| | | | |
Db 49 YLKQHCTAKWTKVACPDPHYT 71

RESULT 3

US-08-239-256-16
; Sequence 16, Application US/08239256
; Patent No. 5585345
; GENERAL INFORMATION:
; APPLICANT: BOIME, IRVING
; APPLICANT: MATZUK, MARTIN M.
; APPLICANT: KEENE, JEFFREY L.
; TITLE OF INVENTION: CTP EXTENDED FORM OF LH
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W.
; CITY: Washington, D.C.
; COUNTRY: USA

; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/239,256
; FILING DATE: 06-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MORASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 29500-20030.12
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-239-256-16

Query Match 35.8%; Score 53; DB 1; Length 38;
Best Local Similarity 52.9%; Pred. No. 4.2;
Matches 9; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 5 PCRRDSPTTCGCPPPRH 21
|| | | | |
Db 14 PCTCDPFCRCGCSCH 30

RESULT 4

PCT-US96-12374-2
; Sequence 2, Application PC/TUS9612374
; GENERAL INFORMATION:
; APPLICANT: Northwestern University
; TITLE OF INVENTION: Herpes Virus Entry Mediator
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Goldsmith, Milnamow & Katz, Ltd.
; STREET: 180 N. Stetson, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/12374
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Northrup, Thomas E.
; REGISTRATION NUMBER: 33,268
; REFERENCE/DOCKET NUMBER: NOR3446P020PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5400
; TELEFAX: (312) 616-5460
; TELEX: --
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 283 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
TELEX: 910 277299

```

; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 416 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   FEATURE:
;     NAME/KEY: Protein
;     LOCATION: 1..416
;     OTHER INFORMATION: /note= "Xaa refers to stop codon in
;       the open reading frame."
;
; US-08-117-083-62
;
; Query Match 33.8%; Score 50; DB 1; Length 416;
; Best Local Similarity 39.1%; Pred. No. 76;
; Matches 9; Conservative 2; Mismatches 12; Indels 0; Gaps 0;
;
; QY 1 FVORPCRDRSPTTCGCPPRHYT 23
; Db 193 FAKVRCNGNDKTCRCPPHTYT 215
;
; RESULT 8
; US-08-449-645A-30
; Sequence 30, Application US/08449645A
; Patent No. 5981245
; GENERAL INFORMATION:
; APPLICANT: Fox, Gary M.
; TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Patent Operations/RBW
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702.367A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-287
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 967 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   US-08-702-367A-30
;
; Query Match 33.8%; Score 50; DB 2; Length 967;
; Best Local Similarity 45.8%; Pred. No. 1.6e+02;
; Matches 11; Conservative 2; Mismatches 5; Indels 6; Gaps 2;
;
; QY 6 CR----RDSP--TTCGCPPRHYT 23
; Db 257 CRPGFFKASPHIQSCGKCPHSHYT 280
;
; RESULT 10
; US-08-449-645A-13
; Sequence 13, Application US/08449645A
; Patent No. 5981245
; GENERAL INFORMATION:
; APPLICANT: Fox, Gary M.
; TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Patent Operations/RBW
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449.645A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-287
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 967 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   US-08-449-645A-30
;
; Query Match 33.8%; Score 50; DB 2; Length 967;
; Best Local Similarity 45.8%; Pred. No. 1.6e+02;
; Matches 11; Conservative 2; Mismatches 5; Indels 6; Gaps 2;
;
; QY 6 CR----RDSP--TTCGCPPRHYT 23
; Db 257 CRPGFFKASPHIQSCGKCPHSHYT 280
;
; RESULT 9
; US-08-702-367A-30
```

```

; Sequence 30, Application US/08702367A
; Patent No. 5981246
; GENERAL INFORMATION:
; APPLICANT: Fox, Gary M.
; TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Patent Operations/RBW
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702.367A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-287
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 967 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   US-08-702-367A-30
;
; Query Match 33.8%; Score 50; DB 2; Length 967;
; Best Local Similarity 45.8%; Pred. No. 1.6e+02;
; Matches 11; Conservative 2; Mismatches 5; Indels 6; Gaps 2;
;
; QY 6 CR----RDSP--TTCGCPPRHYT 23
; Db 257 CRPGFFKASPHIQSCGKCPHSHYT 280
;
; RESULT 10
; US-08-449-645A-13
; Sequence 13, Application US/08449645A
; Patent No. 5981245
; GENERAL INFORMATION:
; APPLICANT: Fox, Gary M.
; TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Patent Operations/RBW
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449.645A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-287
```

; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 991 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-449-645A-13

Query Match 33.8%; Score 50; DB 2; Length 991;
Best Local Similarity 45.8%; Pred. No. 1.6e+02;
Matches 11; Conservative 2; Mismatches 5; Indels 2;

QY 6 CR----RDSPP--TTGCGPCPPRHVT 23
|| : || : || || || ||
DB 281 CRPGFFKASPHIQSCGKCPPHSYT 304

RESULT 11
US-08-702-367A-13
; Sequence 13, Application US/08702367A
; Patent No. 5981246
; GENERAL INFORMATION:

; APPLICANT: Fox, Gary M.
; TITLE OF INVENTION: Eph-Like Receptor Protein Tyrosine
; TITLE OF INVENTION: Kinases
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Patent Operations/RBW
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-287
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 991 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-702-367A-13

Query Match 33.8%; Score 50; DB 2; Length 991;
Best Local Similarity 45.8%; Pred. No. 1.6e+02;
Matches 11; Conservative 2; Mismatches 5; Indels 2;

QY 6 CR----RDSPP--TTGCGPCPPRHVT 23
|| : || : || || || ||
DB 281 CRPGFFKASPHIQSCGKCPPHSYT 304

RESULT 12
PCT-US95-04681-13
; Sequence 13, Application PC/TUS9504681
; GENERAL INFORMATION:

; APPLICANT: Fox, Gary M.
; TITLE OF INVENTION: Eph-Like Receptor Protein Tyrosine
; TITLE OF INVENTION: Kinases
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Amgen Patent Operations/RBW
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04681
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-287
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 991 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-04681-13

Query Match 33.8%; Score 50; DB 4; Length 991;
Best Local Similarity 45.8%; Pred. No. 1.6e+02;
Matches 11; Conservative 2; Mismatches 5; Indels 2;

QY 6 CR----RDSPP--TTGCGPCPPRHVT 23
|| : || : || || || ||
DB 281 CRPGFFKASPHIQSCGKCPPHSYT 304

RESULT 13
US-08-479-722B-2
; Sequence 2, Application US/08479722B
; Patent No. 6074840
; GENERAL INFORMATION:
; APPLICANT: Bonadio, Jeffrey
; APPLICANT: Yin, Wushan
; TITLE OF INVENTION: LATENT TGF(BINDING PROTEIN (LTBP)
; TITLE OF INVENTION: GENES, COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Williams, Morgan & Amerson
; STREET: 7676 Hillmont, Suite 250
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77040

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,722B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US95/02251
; FILING DATE: 21-FEB-1995

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/316,650
; FILING DATE: 30-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/199,780
; FILING DATE: 18-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Fussey, Shelley P.M.

REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: 4100.000500/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (713) 934-7000
TELEFAX: (713) 934-7011
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1833 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-479-722B-2

Query Match 33.8%; Score 50; DB 3; Length 1833;

Best Local Similarity 77.8%; Pred. No. 2.8e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 12 TTCGCPPPR 20
I:| | | | |
Db 591 TSCAPCPPR 599

RESULT 14
PCT-US95-02251-18
Sequence 18, Application PC/TUS9502251
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STIMULATING BONE
TITLE OF INVENTION: CELLS
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: PatentIn Release #1.0, Version
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02251
FILING DATE: CONCURRENTLY HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/316,650
FILING DATE: 30-SEP-1994
CLASSIFICATION:
APPLICATION NUMBER: US 08/199,780
FILING DATE: 18-FEB-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UMIC009P--
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 1833 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-02251-18

Query Match 33.8%; Score 50; DB 4; Length 1833;

Best Local Similarity 77.8%; Pred. No. 2.8e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 12 TTCGCPPPR 20
I:| | | | |
Db 591 TSCAPCPPR 599

RESULT 15
US-08-147-784-2
Sequence 2, Application US/08147784
Patent No. 5821332
GENERAL INFORMATION:
APPLICANT: Godfrey, Wayne
APPLICANT: Buck, David
APPLICANT: Engleman, Edgar G.
TITLE OF INVENTION: RECEPTOR ON THE SURFACE OF ACTIVATED
TITLE OF INVENTION: CD4+ T-CELLS: ACT-4
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Khourie and Crew
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/147,784
FILING DATE: 03-NOV-1993
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 05490A-220
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 277 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-147-784-2

Query Match 33.1%; Score 49; DB 2; Length 277;
Best Local Similarity 60.0%; Pred. No. 68;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 14 CGPCPPPHYT 23
I:| | | | |
Db 125 CAPCPGHFS 134

Search completed: January 30, 2001, 16:47:49
Job time: 625 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)


```
RESULT 11
Q9JLV2          PRELIMINARY;      PRT;   789 AA.
ID  Q9JLV2
AC  Q9P642;
DT  01-OCT-2000 (TREMBlrel. 15, Created)
DT  01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT  01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE  TRP4-ASSOCIATED PROTEIN TAP1.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=INSULINOMA;
RA  Qian F., Philipson L.H.;
RL  Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AF130458; AAF36513.1; -.
SQ  SEQUENCE 789 AA; 90067 MW;  D6DC90BF2D7C80C CRC64;

Query Match      36.5%; Score 50; DB 11; Length 789;
Best Local Similarity 56.2%; Pred. No. 43;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY  3  GNGPTPRAGRAALQK 18
Db  27  GNGRPRFGNILLQUR 42

RESULT 12
O69549          PRELIMINARY;      PRT;   200 AA.
ID  O69549
AC  O69549;
DT  01-AUG-1998 (TREMBlrel. 07, Created)
DT  01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT  01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE  HYPOTHETICAL 21.2 KDA PROTEIN.
GN  MLCB2548.32C.
OS  Mycobacterium leprae.
OC  Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC  Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX  NCBI_TaxID=1769;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Badcock K., Churcher C.M.;
RL  Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN  [2]
RP  SEQUENCE FROM N.A.
RA  Parkhill J., Barrell B.G.; Rajandream M.A.;
RL  Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN  [3]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=93188700; PubMed=8446027;
RA  Eiglmeter K., Honore N., Woods S.A., Caudron B., Cole S.T.;
RT  "Use of an ordered cosmid library to deduce the genomic organization
RT  of Mycobacterium leprae.";
RL  Mol. Microbiol. 7:197-206(1993).
DR  EMBL; AL023093; CAA18819.1; -.
KW  Hypothetical protein.
SQ  SEQUENCE 200 AA; 21249 MW;  397148F302C51EF9 CRC64;

Query Match      36.1%; Score 49.5; DB 2; Length 200;
Best Local Similarity 48.1%; Pred. No. 12;
Matches 13; Conservative 4; Mismatches 7; Indels 3; Gaps 1;

QY  3  GNGPT---PRAGRAALQKLRRLTEL 26
Db  154  GIAPTLCIGTRSGRPALRISGRRLSRL 180
```

```
RESULT 13
Q9P642          PRELIMINARY;      PRT;   107 AA.
ID  Q9P642
AC  Q9P642;
DT  01-OCT-2000 (TREMBlrel. 15, Created)
DT  01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT  01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE  HYPOTHETICAL 11.8 KDA PROTEIN.
GN  B21J21.60.
OS  Neurospora crassa.
OC  Eukaryota; Fungi; Ascomycota; Sordariales; Sordariaceae; Neurospora.
OX  NCBI_TaxID=5141;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA  Nyakatura G., Mewes H.W., Mannhaupt G.;
RL  Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN  [2]
RP  SEQUENCE FROM N.A.
RA  German Neurospora genome project;
RL  Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AL355929; CAB91328.1; -.
KW  Hypothetical protein.
SQ  SEQUENCE 107 AA; 11795 MW;  0E1434D9BE18ED4B CRC64;

Query Match      35.8%; Score 49; DB 3; Length 107;
Best Local Similarity 47.6%; Pred. No. 7.6;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY  1  PEGWGPTPRAGRAALQKLR 21
Db  40  PWGPGTWDARSSPLQMER 60

RESULT 14
P89451          PRELIMINARY;      PRT;   785 AA.
ID  P89451
AC  P89451;
DT  01-MAY-1997 (TREMBlrel. 03, Created)
DT  01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT  01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE  HERPES SIMPLEX VIRUS TYPE 2 (STRAIN HG52), COMPLETE GENOME.
GN  UL28.
OS  Herpes simplex virus (type 2).
OC  Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC  Alphaherpesvirinae; Simplexvirus.
OX  NCBI_TaxID=10310;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=HG52;
RX  MEDLINE=87111457; PubMed=3027242;
RA  McGeoch D.J., Moss H.W., McNab D., Frame M.C.;
RT  "DNA sequence and genetic content of the HindIII 1 region in the short
RT  unique component of the herpes simplex virus type 2 genome;
RT  identification of the gene encoding glycoprotein G, and evolutionary
RT  comparisons.";
RL  J. Gen. Virol. 68:19-38(1987).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=HG52;
RX  MEDLINE=90278430; PubMed=2161906;
RA  Everett R., Fenwick M.;
RT  "Comparative DNA sequence analysis of the host shutoff genes of
RT  different strains of herpes simplex virus: type 2 strain HG52 encodes
RT  a truncated UL41 product.";
RL  J. Gen. Virol. 71:1387-1390(1990).
RN  [3]
RP  SEQUENCE FROM N.A.
RC  STRAIN=HG52;
RX  MEDLINE=92113549; PubMed=1662697;
RA  McGeoch D.J., Cunningham C., McIntyre G., Dolan A.;
```

RT "Comparative sequence analysis of the long repeat regions and
 RT adjoining parts of the long unique regions in the genomes of herpes
 RT simplex viruses types 1 and 2."
 RL J. Gen. Virol. 72:3057-3075(1991).
 RN [4]
 RC SEQUENCE FROM N.A.
 RC STRAIN=HG52;
 RX MEDLINE=92356101; PubMed=1322965;
 RA Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;
 RT "A novel herpes simplex virus gene (UL49A) encodes a putative membrane
 RT protein with counterparts in other herpesviruses."
 RL J. Gen. Virol. 73:2167-2171(1992).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HG52;
 RA Dolan A.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z86099; CAB06753.1; -.
 DR INTERPRO: IPR000501; -.
 DR PFAM: PF01366; PRTF; 1.
 SQ SEQUENCE 785 AA; 85240 MW; 246988E41997DF62 CRC64;

Query Match 35.8%; Score 49; DB 12; Length 785;
 Best Local Similarity 52.2%; Pred. No. 60;
 Matches 12; Conservative 2; Mismatches 7; Indels 2; Gaps 1;
 QY 1 PEGWGTPPRAGRAALQKLRRRL 23
 ||||| ||||| ::|||
 Db 482 PESWGDI--ATRAADVRRRL 502

RESULT 15
 Q9WVL8
 ID Q9WVL8 PRELIMINARY; PRT; 841 AA.
 AC Q9WVL8;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE ZINC FINGER TYPE TRANSCRIPTION FACTOR MZF-3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B57BL/6;
 RA Yamaguchi Y., Zhang X.O., Suda T.;
 RT "Cloning of MZF-1 related gene, murine MZF-3."
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF082568; AAD38426.1; -.
 DR HSSP: P08045; 1ZNF.
 DR INTERPRO: IPR000822; -.
 DR PFAM: PF00096; zf-C2H2; 15.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2; 14.
 KW Zinc-finger; Metal-binding; DNA-binding.
 SQ SEQUENCE 841 AA; 91149 MW; 8AB1E4D2A02C3516 CRC64;

Query Match 35.8%; Score 49; DB 11; Length 841;
 Best Local Similarity 87.5%; Pred. No. 65;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 PEGWGTPP 8
 ||||| |
 Db 113 PEGWGPNP 120

Search completed: January 30, 2001, 16:54:52
 Job time: 908 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:47:44 ; Search time 132.77 seconds
(without alignments)
3.516 Million cell updates/sec

Title: US-09-518-931-2_COPY_239_264

Perfect score: 137

Sequence: 1 PEGWGTPPRAGRAALQLKRLRTEL 26

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA.*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6_COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	137	100.0	300	2	US-08-794-796-2
2	53.5	39.1	774	2	US-08-231-193A-42
3	53.5	39.1	774	2	US-08-486-273A-42
4	53.5	39.1	774	3	US-08-480-474-42
5	53.5	39.1	774	3	US-08-940-086A-42
6	53.5	39.1	1214	2	US-08-231-193A-54
7	53.5	39.1	1214	2	US-08-486-273A-54
8	53.5	39.1	1214	3	US-08-480-474-54
9	53.5	39.1	1214	3	US-08-940-086A-54
10	53.5	39.1	1219	2	US-08-231-193A-50
11	53.5	39.1	1219	2	US-08-486-273A-50
12	53.5	39.1	1219	3	US-08-480-474-50
13	53.5	39.1	1219	3	US-08-940-086A-50
14	53.5	39.1	1231	2	US-08-231-193A-48
15	53.5	39.1	1231	2	US-08-486-273A-48
16	53.5	39.1	1231	3	US-08-480-474-48
17	53.5	39.1	1231	3	US-08-940-086A-48
18	53.5	39.1	1236	2	US-08-231-193A-6
19	53.5	39.1	1236	2	US-08-486-273A-6
20	53.5	39.1	1236	3	US-08-480-474-6
21	53.5	39.1	1236	3	US-08-940-086A-6
22	53.5	39.1	1239	2	US-08-231-193A-52
23	53.5	39.1	1239	2	US-08-486-273A-52
24	53.5	39.1	1239	3	US-08-480-474-52
25	53.5	39.1	1239	3	US-08-940-086A-52
26	53.5	39.1	1244	2	US-08-231-193A-46
27	53.5	39.1	1244	2	US-08-486-273A-46
28	53.5	39.1	1244	3	US-08-480-474-46

29	53.5	39.1	1244	3	US-08-940-086A-46	Sequence 46, Appl
30	50	36.5	797	1	US-08-698-551-18	Sequence 18, Appl
31	50	36.5	797	2	US-08-839-032A-18	Sequence 18, Appl
32	46	33.6	799	3	US-08-909-95A-4	Sequence 4, Appli
33	45.5	33.2	380	1	US-08-416-478A-6	Sequence 6, Appli
34	45.5	33.2	380	2	US-08-474-988B-6	Sequence 6, Appli
35	45.5	33.2	380	2	US-08-394-442B-6	Sequence 7, Appli
36	45.5	33.2	470	1	US-08-416-478A-7	Sequence 7, Appli
37	45.5	33.2	470	2	US-08-474-988B-7	Sequence 7, Appli
38	45.5	33.2	470	2	US-08-394-442B-7	Sequence 7, Appli
39	45.5	33.2	471	1	US-08-416-478A-2	Sequence 2, Appli
40	45.5	33.2	471	2	US-08-474-988B-2	Sequence 2, Appli
41	45.5	33.2	471	2	US-08-394-442B-2	Sequence 2, Appli
42	45.5	33.2	476	2	US-08-737-271-1	Sequence 1, Appli
43	45.5	33.2	476	3	US-09-058-555-1	Sequence 1, Appli
44	45.5	33.2	498	1	US-08-416-478A-9	Sequence 9, Appli
45	45.5	33.2	498	2	US-08-474-988B-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1
US-08-794-796-2
; Sequence 2, Application US/08794796
; Patent No. 5885800
; GENERAL INFORMATION:
; APPLICANT: Emery, John
; APPLICANT: Tan, KB
; APPLICANT: Trunch, Alem
; APPLICANT: Young, Peter
; TITLE OF INVENTION: Tumor Necrosis Related Receptor,
; TITLE OF INVENTION: TR4
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithkline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/794,796
; FILING DATE: 04-FEB-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: GH50000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-4026
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 300 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-794-796-2

Query Match 100.0%; Score 137; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 1.7e-13;

```

; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR SUBUNITS, DNA
; TITLE OF INVENTION: ENCODING SAME AND USES THEREFOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,273A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/231,193
; FILING DATE: 20-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-9383B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-238-0999
; TELEFAX: 619-238-0062
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 774 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-486-273A-42

Query Match          39.1%; Score 53.5; DB 2; Length 774;
Best Local Similarity 73.3%; Pred. No. 2.7;
Matches 11; Conservative 0; Mismatches 3; Indels 1; Gaps

Qy      1 PEGWGPTRAGRAAL 15
        | |||| | |||||
Db      692 PTGWGP-PDGGRAAL 705

RESULT      4
US-08-480-474-42
; Sequence 42, Application US/08480474
; Patent No. 6033865
; GENERAL INFORMATION:
; APPLICANT: Daggett, Lorrie P.
; APPLICANT: Ellis, Steven B.
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lu, Chin-Chun
; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR SUBUNITS, DNA
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/480,474
FILING DATE: 06-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-9382B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-238-0999
TELEFAX: 619-238-0062
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 774 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-474-42

Query Match 39.1%; Score 53.5; DB 3; Length 774;
Best Local Similarity 73.3%; Pred. No. 2.7;
Matches 11; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 PEGWGTPPRAGRAAL 15
| | | | | | | | | |
Db 692 PTGWGP-PDGGRAAL 705

RESULT 5
US-08-940-086A-42
; Sequence 42, Application US/08940086A
; Patent No. 6111091
; GENERAL INFORMATION:
; APPLICANT: Daggett, Lorrie P.
; APPLICANT: Ellis, Steven B.
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lu, Chin-Chun
; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McAuliffe
; STREET: 4250 Executive Square, 7th Floor
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/940,086A
; FILING DATE: 29-SEPT-97
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/231,193
; FILING DATE: 20-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/052,449
; FILING DATE: 20-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24735-9383C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 450-8400
; TELEFAX: (619) 450-8499
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 774 amino acids
; TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-940-086A-42

Query Match 39.1%; Score 53.5; DB 3; Length 774;
Best Local Similarity 73.3%; Pred. No. 2.7;
Matches 11; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 PEGWGTPPRAGRAAL 15
| | | | | | | | | |
Db 692 PTGWGP-PDGGRAAL 705

RESULT 6
US-08-231-193A-54
; Sequence 54, Application US/08231193A
; Patent No. 5849895
; GENERAL INFORMATION:
; APPLICANT: Daggett, Lorrie P.
; APPLICANT: Ellis, Steven B.
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lu, Chin-Chun
; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,193A
; FILING DATE: 20-APR-1994
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/052,459
; FILING DATE: 20-APR-1993
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-9383
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-238-0999
; TELEFAX: 619-238-0062
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1214 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-231-193A-54

Query Match 39.1%; Score 53.5; DB 2; Length 1214;
Best Local Similarity 73.3%; Pred. No. 4.5;
Matches 11; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 PEGWGTPPRAGRAAL 15
| | | | | | | | | |
Db 935 PTGWGP-PDGGRAAL 948

RESULT 7
US-08-486-273A-54

```

; Sequence 54, Application US/08486273A
; Patent No. 5985586
;
; GENERAL INFORMATION:
; APPLICANT: Daggett, Lorrie P.
; APPLICANT: Ellis, Steven B.
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lu, Chin-Chun
; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR SUBUNITS, DNA
; TITLE OF INVENTION: ENCODING SAME AND USES THEREFOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92101-2926
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,273A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/231,193
; FILING DATE: 20-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-9383B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-238-0099
; TELEFAX: 619-238-0062
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1214 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-08-486-273A-54
;
; Query Match 39.1%; Score 53.5; DB 2; Length 1214;
; Best Local Similarity 73.3%; Pred. No. 4-5;
; Matches 11; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
;
Qy 1 PEGMGPTPRGRAAL 15
Db 935 PTWGWP-PDGGRAAL 948
;
;
RESULT 8
US-08-480-474-54
; Sequence 54, Application US/08480474
; Patent No. 6033865
;
; GENERAL INFORMATION:
; APPLICANT: Daggett, Lorrie P.
; APPLICANT: Ellis, Steven B.
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lu, Chin-Chun
; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR SUBUNITS, DNA ENCODING
; TITLE OF INVENTION: SAME AND USES THEREFOR
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: U.S.A.
;

```

```

;
;
; TELECOMMUNICATION INFORMATION:
;
; TELEPHONE: (619) 450-8400
;
; TELEFAX: (619) 450-8499
;
; INFORMATION FOR SEQ ID NO: 54:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 1214 amino acids
;
; TYPE: amino acid
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: protein
;
; US-08-940-086A-54

```

```
Query Match          39.1%; Score 53.5; DB 3; Length 1214;
Best Local Similarity 73.3%; Pred. NO. 4.5;
Matches 11; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
```

Qy	1	PEGWGPTPRAGRAAL	15
D _b	935	PTGWGP-PDGGRAAL	948

RESULT 10
US-08-231-193A-50
; Sequence 50, Application US/08231193A
; Patent No. 5849895
; GENERAL INFORMATION:
; APPLICANT: Daggett, Lorrie P.
; APPLICANT: Ellis, Steven B.
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lu, Chin-Chun
; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR
; TITLE OF INVENTION: SUBUNITS, NUCLEIC ACIDS ENCODING SAME AND USES THEREFOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
;

Query Match 39.1%; Score 53.5; DB 2; Length 1219;
Best Local Similarity 73.3%; Pred. No. 4.5;
Matches 11; Conservative 0; Mismatches 3; Indels 1

Qy	1	PEGWGPTPRAGRAAL	15
Db	940	PTWGWP-PDGGRAAL	953
RESULT 11			
US-08-486-273A-50			
: Sequence 50, Application US/08486273A			
: Patent No. 5985586			
: GENERAL INFORMATION:			

Query Match	39.1%;	Score 53.5;	DB 2;	Length 1219;
Best Local Similarity	73.3%;	Pred. No. 4.5;		
Matches 11;	Conservative	0;	Mismatches 3;	Indels 1;
				Gaps 1;

Qy 1 PEGWGPTPRAGRAAL 15
| | | | |
Db 940 PTGWGPP-PDGGRAAL 953

```

RESULT 12
US-08-480-474-50
; Sequence 50, Application US/08480474
; Patent No. 6033865
; GENERAL INFORMATION:
; APPLICANT: Daggett, Lorrie P.
; APPLICANT: Ellis, Steven B.
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lu, Chin-Chun
; TITLE OF INVENTION: HUMAN N-METH
; TITLE OF INVENTION: SAME AND USE

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/052,449
; FILING DATE: 20-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24735-9383C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 450-8400
; TELEFAX: (619) 450-8499
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1219 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-940-086A-50

Query Match 39.1%; Score 53.5; DB 3; Length 1219;
Best Local Similarity 73.3%; Pred. No. 4.5;
Matches 11; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 PEGWGPTPRAGRAAL 15
Db 940 PTGWGP-PDGGRAAL 953

RESULT 14
US-08-231-193A-48
; Sequence 48, Application US/08231193A
; Patent No. 5649895
; GENERAL INFORMATION:
; APPLICANT: Daggett, Lorrie P.
; APPLICANT: Ellis, Steven B.
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lu, Chin-Chun
; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR
; TITLE OF INVENTION: SUBUNITS, NUCLEIC ACIDS ENCODING SAME AND USES THEREFOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,193A
; FILING DATE: 20-APR-1994
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/052,459
; FILING DATE: 20-APR-1993
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-9383
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-238-0999
; TELEFAX: 619-238-0062
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1231 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

```


US-08-231-193A-48

Query Match 39.1%; Score 53.5; DB 2; Length 1231;
Best Local Similarity 73.3%; Pred. No. 4.6;
Matches 11; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 PEGWGPTPRAGRAAL 15
| | | | | | | | | |
Db 952 PTCWGP-PDGGRAAL 965

RESULT 15

US-08-486-273A-48
; Sequence 48, Application US/08486273A
; Patent No. 5985586
; GENERAL INFORMATION:
; APPLICANT: Daggett, Lorrie P.
; APPLICANT: Ellis, Steven B.
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lu, Chin-Chun
; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR SUBUNITS, DNA
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,273A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/231,193
; FILING DATE: 20-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-9383B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-238-0999
; TELEFAX: 619-238-0062
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1231 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-486-273A-48

Query Match 39.1%; Score 53.5; DB 2; Length 1231;
Best Local Similarity 73.3%; Pred. No. 4.6;
Matches 11; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 PEGWGPTPRAGRAAL 15
| | | | | | | | | |
Db 952 PTCWGP-PDGGRAAL 965

Search completed: January 30, 2001, 16:47:45
Job time: 621 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:45:25 ; Search time 183.67 Seconds
(without alignments)
2.979 Million cell updates/sec

Title: US-09-518-931-2_COPY_283_298

Perfect score: 81

Sequence: 1 ARMPGLERSVREFLP 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq_36: *
- 1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT: *
 - 2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT: *
 - 3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT: *
 - 4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT: *
 - 5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT: *
 - 6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT: *
 - 7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT: *
 - 8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT: *
 - 9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT: *
 - 10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT: *
 - 11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT: *
 - 12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT: *
 - 13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT: *
 - 14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT: *
 - 15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT: *
 - 16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT: *
 - 17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT: *
 - 18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT: *
 - 19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT: *
 - 20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT: *
 - 21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	215	20 W93585	Human hAPO6 protei
2	81	100.0	271	20 Y42184	Human mFLINT #1 pr
3	81	100.0	273	20 Y42185	Human mFLINT #2 pr
4	81	100.0	300	19 W66102	Amino acid sequenc
5	81	100.0	300	19 W63622	Human tumour necro
6	81	100.0	300	20 Y03099	Human lung TNF-rec
7	81	100.0	300	20 Y42182	Human FLINT #1 pro
8	81	100.0	300	20 Y17479	Mammalian tumour n
9	81	100.0	300	20 Y06817	Human DCR3 polypep
10	81	100.0	300	20 W97749	Human tumour necro
11	81	100.0	300	20 W95082	Orphan receptor (H
12	81	100.0	300	21 Y77458	Human TNF receptor

13	81	100.0	302	20 Y42183	Human FLINT #2 pro
14	42	51.9	758	21 Y84408	Amino acid sequenc
15	41	50.6	911	19 W74723	Human secreted pro
16	40	49.4	936	17 W06719	Rhodopseudomonas s
17	39	48.1	350	16 R76598	Deduced human prot
18	38.5	47.5	331	21 Y99372	Human PRO1430 (UNQ
19	38	46.9	343	19 W60053	Human protein tyro
20	38	46.9	400	18 W17831	Human mavalonate p
21	38	46.9	839	13 R22275	Bovine liver GPI-p
22	38	46.9	839	16 R75109	Glycosyl-phosphati
23	37	45.7	113	20 Y13166	Human secreted pro
24	37	45.7	369	19 W62647	Mature durum wheat
25	37	45.7	420	20 Y13362	Amino acid sequenc
26	37	45.7	420	20 Y05281	EGF-like homologue
27	37	45.7	420	21 Y88569	Human PRO214 amino
28	37	45.7	434	21 Y76151	Human secreted pro
29	36.5	45.1	539	20 Y35125	Chlamydia pneumoni
30	36.5	45.1	376	14 R36773	Human fibromodulin
31	36.5	45.1	376	18 W26404	Human fibromodulin
32	36	44.4	14	16 W21527	Prorenin derived s
33	36	44.4	16	21 Y79294	Human prorenin pro
34	36	44.4	39	7 F60296	Sequence attached
35	36	44.4	43	21 Y79293	Human prorenin pro
36	36	44.4	44	7 P61070	Peptide antigenic f
37	36	44.4	75	20 Y02704	Human secreted pro
38	36	44.4	224	14 R52988	Canine hookworm Ne
39	36	44.4	224	20 Y23596	Canine hookworm ne
40	36	44.4	263	20 Y23602	Canine hookworm ne
41	36	44.4	270	20 Y23608	Canine hookworm ne
42	36	44.4	274	14 R42488	Canine hookworm Ne
43	36	44.4	274	20 Y23600	Canine hookworm ne
44	36	44.4	274	20 Y23592	Canine hookworm ne
45	36	44.4	274	20 Y23591	Neutrophil inhibit

ALIGNMENTS

RESULT 1

W93585
ID W93585 standard; Protein; 215 AA.

XX W93585;

XX 18-JUN-1999 (first entry)

XX Human hAPO6 protein.

XX Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;

XX developmental abnormality; gestational abnormality; prostate cancer;

XX APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;

XX cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;

XX apoptosis; human.

XX Homo sapiens.

XX OS

XX W09911791-A2.

XX 11-MAR-1999.

XX 04-SEP-1998; 98WO-US18393.

XX 05-SEP-1997; 97US-0924634.

XX (UNIW) UNIV WASHINGTON.

XX Chaudhary PM;

XX WPI; 1999-205191/17.

XX N-PSDB; X23419.

XX New Tumor Necrosis Factor family receptor polypeptides and ligands -

XX useful for diagnosis and treatment of prostate cancer and

PR 18-DEC-1998; 98US-0112933.
 PR 22-DEC-1998; 98US-0113407.
 XX
 PA (ELIL) LILLY & CO ELI.
 XX
 PI Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
 PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
 PI Song HW, Wang J, Wu X, Zuckerman SH;
 XX
 DR WPI: 1999-591319/50.
 DR N-PSDB: Z25378.
 XX
 PT Use of mature FLINT for treating acute liver failure, inflammation,
 PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
 PT and proinflammatory activity
 XX
 PS Example 2: Fig 4; 99pp; English.
 XX
 CC The present invention describes therapeutic applications of mature FLINT
 CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
 CC (mFLINT), which is a member of the tumour necrosis factor receptor
 CC superfamily, is used for treating acute liver failure, inflammation of
 CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
 CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
 CC injury or disorder such as hypercoagulation (including use with
 CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
 CC or disorder, Type I diabetes, cancer, cell damage or damage to an
 CC innocent bystander tissue that is induced by a chemotherapeutic agent or
 CC therapeutic irradiation, treating haematopoietic progenitor cells that
 CC have been exposed to therapeutic radiation or chemotherapy, aplastic
 CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
 CC also used for promoting the growth or differentiation of a haematopoietic
 CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
 CC resulting from abnormal myocardial ischaemia. The present sequence
 CC represents human mFLINT.
 XX
 SQ Sequence 273 AA;

Query Match 100.0%; Score 81; DB 20; Length 273;
 Best Local Similarity 100.0%; Pred. No. 2.6e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ARMPGLSVSRERFLP 16
 |||||
 Db 256 armpglsvsrerflp 271

RESULT 4
 W66102
 ID W66102 standard; Protein: 300 AA.
 XX
 AC W66102;

DT 02-DEC-1998 (first entry)

DE Amino acid sequence of tumour necrosis related receptor (TR4).

KW Human; tumour necrosis related receptor; TR4; agonist; antagonist;
 KW inhibition; chronic; acute; inflammation; arthritis; septicaemia;
 KW autoimmune disease; transplant rejection; stroke; cancer;
 KW Alzheimer's disease.

XX Homo sapiens.

OS EP861850-A1.

PN 02-SEP-1998.

XX 20-JAN-1998; 98EP-0300382.

PR 04-FEB-1997; 97US-0794796.

XX

PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX Emery J, Tan KB, Truneh A, Young PR;
 XX
 DR WPI: 1998-508248/44.
 DR N-PSDB: V07654.
 XX
 PT New DNA encoding tumour necrosis related receptor - used to treat
 PT and prevent e.g. inflammation, arthritis, septicaemia, autoimmune
 PT diseases, transplant rejection, infection, stroke, ischaemia, ARDS,
 PT restenosis, AIDS, bone disorders and cancer
 XX
 PS Claim 1; Fig 1; 21pp; English.
 XX
 CC This is the amino acid sequence of the human tumour necrosis related
 CC receptor (TR4), used in the method of the invention. The TR4 protein
 CC or its agonist can be used to treat a subject in need of enhanced
 CC TR4 polypeptide activity. The antagonist is used to inhibit TR4
 CC polypeptide activity. The active agents can be used for the
 CC treatment and prevention of diseases such as chronic and acute
 CC inflammation, arthritis, septicaemia, autoimmune diseases, transplant
 CC rejection, stroke, cancer, Alzheimer's disease.
 XX
 SQ Sequence 300 AA;

Query Match 100.0%; Score 81; DB 19; Length 300;
 Best Local Similarity 100.0%; Pred. No. 2.9e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ARMPGLSVSRERFLP 16
 |||||
 Db 283 armpglsvsrerflp 298

RESULT 5

W63622
 ID W63622 standard; Protein: 300 AA.

XX W63622;

DT 26-OCT-1998 (first entry)

DE Human tumour necrosis factor receptor-6 alpha protein.

XX Human tumour necrosis factor receptor-6 alpha; TNFR-6 beta;
 KW endothelial cells; keratinocytes; normal prostate; apoptosis;
 KW prostate tumour tissue.

XX Homo sapiens.

FH Key Location/Qualifiers

FT Peptide 1..30

FT Protein 31..300 /note= "TNFR-6 alpha"

FT Region 31..282

FT /note= "Soluble extracellular domain"

XX WO9830694-A2.

PN 16-JUL-1998.

XX 13-JAN-1998; 98WO-US00153.

PR 14-JAN-1997; 97US-0035496.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Ebner R, Feng P, Gentz RL, Ni J, Ruben SM, Yu G;

XX WPI: 1998-399142/34.

DR N-PSDB: V39085.

XX

PT Human tumour necrosis factor receptors 6-alpha and 6-beta - used in
 PT the diagnosis of immune system-related disorder(s)

XX
 PS
 CC Claim 20; Fig 1; 91pp; English.

XX The present sequence represents the human tumour necrosis factor
 CC receptor-6 alpha (TNFR-6 alpha) protein. The invention also provides
 CC for the TNFR-6 beta protein (W63523). TNFR-6 alpha and TNFR-6 beta are
 CC members of the tumour necrosis factor receptor (TNFR) family. TNFRs
 CC are expressed in endothelial cells, keratinocytes, normal prostate and
 CC prostate tumour tissue. For a number of disorders of these cells,
 CC particularly of the immune system, substantially altered (whether
 CC increased or decreased) levels of TNFR-6 alpha and/or TNFR-6 beta
 CC expression can be detected, therefore the TNFR-6 alpha and TNFR-6 beta
 CC polypeptides, nucleic acids and antibodies are claimed to be useful in
 CC the diagnosis of such disorders. Mutations of the TNFR-6 alpha and
 CC TNFR-6 beta genes can also be detected. The TNFR polypeptides are
 CC also claimed to be useful for identifying ligands which may be useful
 CC in the treatment of apoptosis related disorders.

XX
 SQ Sequence 300 AA;

Query Match 100.0%; Score 81; DB 19; Length 300;
 Best Local Similarity 100.0%; Pred. No. 2.9e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARMPGLRSVRERFLP 16
 |||||

Db 283 armpglrsvrerflp 298
 |||||

RESULT 6

Y03099
 ID Y03099 standard; Protein; 300 AA.

XX
 AC Y03099;

XX
 DT 09-DEC-1999 (first entry)

XX Human lung TNF-receptor protein.

XX Tumour necrosis factor; TNF; TNF receptor; human; lung; gene therapy;
 KW detection; immunoassay; diagnosis; disease; immune system; tumour;
 KW osteogenic system; cardiovascular system; central nervous system; asthma;
 KW peripheral nervous systems; transplant incompatibility; antitumor;
 KW rheumatoid arthritis; antiasthmatic; antiarthritic.

XX Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 134..1036
 FT /*tag= a
 FT /product= "TNF-receptor"

XX DE19809978-A1.

XX 16-SEP-1999.

XX 09-MAR-1998; 98DE-1009978.

XX 09-MAR-1998; 98DE-1009978.

XX (BADI) BASE AG.

XX Kroeger B;

XX WPI; 1999-519473/44.

XX N-PSDB; Z09998.

XX New soluble member of tumor necrosis factor receptor family, useful for
 PT identification specific modulators and for treating disease e.g. tumors

XX
 PS
 CC Claim 1; Page 8-9; 10pp; German.

XX This invention describes a novel tumour necrosis factor (TNF) receptor
 CC (I) isolated from human lung tissue. (I) is used: (i) to raise specific
 CC antibodies (Ab); (ii) to screen for specific (ant)agonists or ligands
 CC (A), potential therapeutic agents; and (iii) therapeutically (optionally
 CC expressed from a gene therapy vector) in conditions associated with a
 CC deficit of (I). Ab are used: (a) for qualitative or quantitative
 CC detection of (I) in standard immunoassays (for diagnosis of disease, or
 CC susceptibility, or for monitoring); and (b) as therapeutic inhibitors in
 CC cases where (I) is overexpressed. Nucleic acid (II) that encodes (I) is
 CC used: (A) for recombinant production of (I); (B) also its oligonucleotide
 CC fragments, in standard hybridization and/or amplification assays; (C) as
 CC source of antisense molecules or ribozymes; and (D) to produce transgenic
 CC animals (for studying (patho)physiology of (I)). Diseases possibly
 CC associated with under- or over-expression of (I) are those of the immune,
 CC osteogenic, cardiovascular and central or peripheral nervous systems,
 CC tumors, transplant incompatibility, asthma and rheumatoid arthritis. The
 CC products of the invention have antitumor, antiasthmatic and
 CC antiarthritic activity. This sequence represents the TNF-receptor of the
 CC invention.

XX
 SQ Sequence 300 AA;

Query Match 100.0%; Score 81; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 2.9e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARMPGLRSVRERFLP 16
 |||||

Db 283 armpglrsvrerflp 298
 |||||

RESULT 7

Y42182
 ID Y42182 standard; Protein; 300 AA.

XX
 AC Y42182;

XX
 DT 17-DEC-1999 (first entry)

XX Human FLINT #1 protein sequence.

XX Human; FLINT; mFLINT; OPC3; tumour necrosis factor receptor; FasL;
 KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
 KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
 KW reperfusion-associated injury; aplastic anaemia; differentiation;
 KW growth; myelodysplastic syndrome; pancytopenic condition;
 KW myocardial ischaemia.

XX Homo sapiens.

XX W09950413-A2.

XX 07-OCT-1999.

XX 30-MAR-1999; 99WO-US06797.

XX 30-MAR-1998; 98US-0079856.

XX 20-MAY-1998; 98US-0086074.

XX 09-SEP-1998; 98US-0099643.

XX 17-DEC-1998; 98US-0112577.

XX 18-DEC-1998; 98US-0112703.

XX 18-DEC-1998; 98US-0112933.

XX 22-DEC-1998; 98US-0113407.

XX (ELIL) LILLY & CO ELI.

XX Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
 PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
 PI Song HY, Wang J, Wu X, Zuckerman SH;

XX WPI: 1999-591319/50.
 DR N-PSDB; Z25375.
 XX
 PT Use of mature FLINT for treating acute liver failure, inflammation,
 PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
 PT and proinflammatory activity
 XX
 PS Claim 30; Fig 1; 99pp; English.
 XX
 CC The present invention describes therapeutic applications of mature FLINT
 CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
 CC (mFLINT), which is a member of the tumour necrosis factor receptor
 CC superfamily, is used for treating acute liver failure, inflammation of
 CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
 CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
 CC injury or disorder such as hypercoagulation (including use with
 CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
 CC or disorder, Type I diabetes, cancer, cell damage or damage to an
 CC innocent bystander tissue that is induced by a chemotherapeutic agent or
 CC therapeutic irradiation, treating haematopoietic progenitor cells that
 CC have been exposed to therapeutic radiation or chemotherapy, aplastic
 CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
 CC also used for promoting the growth or differentiation of a haematopoietic
 CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
 CC resulting from abnormal myocardial ischaemia. The present sequence
 CC represents human FLINT.
 XX
 SQ Sequence 300 AA;
 Query Match 100.0%; Score 81; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 2.9e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ARMPGLSVRERFLP 16
 Db 283 armpglsvrflp 298
 |||||
 RESULT 8
 Y17479
 ID Y17479 standard; Protein; 300 AA.
 AC Y17479;
 XX
 XX 02-AUG-1999 (first entry)
 DT
 XX Mammalian tumour necrosis factor receptor OPG-2.
 DE
 XX Tumour necrosis factor receptor; TNF receptor; OPG-2; Paget's disease;
 KW osteopenic disorder; osteoclast activity; primary osteoporosis;
 KW hyperglycaemia; osteolytic metastasis; immune response; cancer.
 XX
 OS Mammalia.
 XX
 PN WO9926977-A1.
 XX
 XX 03-JUN-1999.
 PD
 XX 24-NOV-1998; 98WO-US25065.
 PF
 XX 17-FEB-1998; 98US-0074896.
 PR
 XX 24-NOV-1997; 97US-0066446.
 PR
 XX (BIOJ) BIOGEN INC.
 PA
 XX Tschoep J;
 PI
 DR WPI: 1999-347693/29.
 DR N-PSDB; X76052.
 XX
 XX New tumour necrosis factor family receptor OPG-2
 PT

XX Claim 1; Page 18; 22pp; English.
 PS
 XX
 CC The present sequence represents a mammalian tumour necrosis factor
 CC receptor, designated OPG-2. OPG-2, is a member of the tumour necrosis
 CC factor receptor family, and can be used: (i) to raise specific
 CC antibodies (Ab), (ii) to treat osteopenic disorders associated with
 CC excessive osteoclast activity, e.g. primary osteoporosis, Paget's
 CC disease, hyperglycaemia of malignancy, or osteolytic metastases; (iii)
 CC for affinity purification of cognate ligands, and (iv) to screen for
 CC ligands (antagonists or agonists). Ab, or other OPG-2 blocking agents
 CC such as soluble forms of the protein, are used to prevent, or reduce
 CC severity of, an immune response, and for treating cancer. They can also
 CC be used in diagnostic assays. The nucleic acid sequence encoding OPG-2
 CC can be used as a probe to isolate related sequences from other species.
 XX
 SQ Sequence 300 AA;
 Query Match 100.0%; Score 81; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 2.9e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ARMPGLSVRERFLP 16
 Db 283 armpglsvrflp 298
 |||||
 RESULT 9
 Y06817
 ID Y06817 standard; Protein; 300 AA.
 XX
 AC Y06817;
 XX
 XX 24-JUN-1999 (first entry)
 DT
 XX Human Dcr3 polypeptide.
 DE
 XX Dcr3 polypeptide; tumour necrosis factor receptor; TNFR; Fas ligand;
 KW apoptosis; T cell mediated immune response; allergy; asthma; cancer;
 KW rheumatoid arthritis; Crohn's disease; guest vs. host disease; human;
 KW gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO9914330-A1.
 XX
 XX 25-MAR-1999.
 PD
 XX 18-SEP-1998; 98WO-US19661.
 PF
 XX 30-JUL-1998; 98US-0094640.
 PR
 XX 18-SEP-1997; 97US-0059288.
 PR
 XX (GETH) GENENTECH INC.
 PA
 XX Ashkenazi AJ, Botstein D, Dodge KH, Goddard A, Gurney AL;
 PI Kim KJ, Lawrence DA, Pitti R, Roy MA, Tumas DB;
 PI Wood WI;
 XX
 XX WPI: 1999-244032/20.
 DR N-PSDB; X32744.
 DR
 XX Dcr3 polypeptide related to tumor necrosis factor receptor
 PT
 XX Claim 5; Fig 1; 88pp; English.
 PS
 XX This represents a human Dcr3 polypeptide, a homologue of tumour necrosis
 CC factor receptor (TNFR) polypeptide. Host cells containing a vector
 CC comprising the Dcr3 nucleic acid can be used for the recombinant
 CC expression of the protein. Dcr3 binds to Fas ligand, so it (or its
 CC chimeras) are useful for modulating apoptosis in mammalian cells, also
 CC other Fas-ligand induced activities, particularly to inhibit T cell

CC mediated immune responses, e.g. in treatment of allergy, asthma,
CC rheumatoid arthritis, Crohn's disease, guest vs. host disease etc. Dcr3
CC may also be used to identify specific binding proteins, potential
CC inhibitors. Antibodies against Dcr3 are used to treat cancer,
CC specifically of the lung and colon, also in diagnosis and for affinity
CC purification of the protein. Detecting mutations in the gene for Dcr3 is
CC also used to diagnose cancer, or predisposition to it. Dcr3 nucleic acid
CC is useful as hybridization probe to detect genomic or related sequences;
CC for chromosome and gene mapping; as source of antisense sequences; for
CC expression of recombinant Dcr3 and to generate transgenic animals (for
CC development and screening of therapeutic agents), also for in vivo or
CC ex vivo gene therapy.
XX
SQ Sequence 300 AA;

Query Match 100.0%; Score 81; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.9e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARMPGLERSVRERFLP 16
DB 283 armpglersvrerflp 298
|||||

RESULT 10
ID W97749 standard; Protein: 300 AA.
XX AC W97749;
XX DT 21-MAY-1999 (first entry)
XX DE Human tumour necrosis factor receptor ZTNFR-5.
XX KW ZTNFR-5; tumour necrosis factor receptor; TNFR; human;
XX KW cell maturation; bone cell regulation.
XX OS Homo sapiens.

FT Key Location/Qualifiers
FT FH 1..23 /note= "signal peptide"
FT FT 24..300 /note= "mature protein"
FT FT 24..194 /note= "extracellular domain"
FT FT 49..71 /note= "cysteine-rich pseudo-repeat 1"
FT FT 72..113 /note= "cysteine-rich pseudo-repeat 1"
FT FT 114..151 /note= "cysteine-rich pseudo-repeat 1"
FT FT 152..194 /note= "cysteine-rich pseudo-repeat 1"

XX WO9904001-A1.
XX PD 28-JAN-1999.
XX PF 21-JUL-1998; 98WO-US15072.
XX PR 21-JUL-1997; 97US-0053203.
XX PA (ZYMO) ZYMOGENETICS INC.
XX PI Farrah TM;
XX DR WPI; 1999-132245/11.
XX DR N-PSDB; X07226.
XX PT Novel tumour necrosis factor receptor ZTNFR5 - useful for
XX regulating maturation of TNF-ligand bearing cells

XX
PS Claim 1; Page 84-85; 109pp; English.
XX This polypeptide comprises a new, secreted tumour necrosis factor
CC receptor (see W97749), designated ZTNFR-5. Novel ZTNFR-5 encoding
CC polynucleotides and polypeptides were initially identified by
CC querying an expressed sequence tag (EST) database for sequences
CC homologous to conserved motifs within the TNF receptor family.
CC Based on this search, a contig of 16 ESTs (see X07226) was
CC constructed. ZTNFR-5 polypeptides comprise 4 cysteine-rich repeats
CC (see also W97750-55) that are homologous to other TNF receptors, in
CC particular the soluble, secreted TNF receptor osteoprotegerin.
CC ZTNFR-5 polypeptide can be prepared by recombinant methods. The
CC polypeptide, especially the extracellular domain, can be used to
CC generate a soluble variant of ZTNFR-5. The polypeptides and
CC nucleic acids can be used to screen for ligands, agonists and
CC antagonists of ZTNFR-5. The polypeptides can be used in bone cell
CC regulation and to regulate the maturation of TNF ligand-bearing
CC cells such as T- or B-cells, lymphocytes, peripheral blood
CC mononuclear cells, polymorphonuclear leukocytes, fibroblasts or
CC haematopoietic cells.
XX
SQ Sequence 300 AA;

Query Match 100.0%; Score 81; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.9e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARMPGLERSVRERFLP 16
DB 283 armpglersvrerflp 298
|||||

RESULT 11
W95082
ID W95082 standard; Protein: 300 AA.
XX AC W95082;
XX DT 20-MAY-1999 (first entry)
XX DE Orphan receptor (HUMAN NTR-1) polypeptide.
XX KW HUMAN NTR-1; orphan receptor; osteoprotegerin; OPG; TNFR; human;
XX KW tumour necrosis factor receptor; muscle disorder; bone mass; screening;
XX KW muscle metabolism; binding agent; cognate ligand.
XX OS Homo sapiens.
XX PN WO9907738-A2.
XX PD 18-FEB-1999.
XX PF 04-AUG-1998; 98WO-US16202.
XX PR 06-AUG-1997; 97US-0054869.
XX PA (PROC) PROCTER & GAMBLE CO.
XX PA (REGE-) REGENERON PHARM INC.
XX PI Masiakowski PJ, Morris J, Valenzuela DM;
XX WPI; 1999-167365/14.
XX DR N-PSDB; X22300.
XX PT Novel orphan human receptor polypeptide and nucleic acid - useful as
XX diagnostic reagents and for treatment of muscle disorders
XX Claim 7; Page 21; 23pp; English.
XX This represents a HUMAN NTR-1 polypeptide, a novel orphan receptor. The
CC protein is related to osteoprotegerin (OPG) and to tumour necrosis factor

CC receptor (TNFR). Host cells transformed with a vector comprising the
 CC HUMAN NTR-1 nucleic acid are used for the recombinant expression of the
 CC protein. HUMAN NTR-1 proteins and antibodies immuno specific for the
 CC protein are useful for diagnosis and treatment of humans and animals,
 CC especially muscle disorders, as the receptor is involved in regulation of
 CC bone mass and muscle metabolism. HUMAN NTR-1 receptors are also useful
 CC for screening for novel binding agents, and cognate ligands, which may be
 CC used to treat disorders associated with HUMAN NTR-1 imbalance.

XX Sequence 300 AA;

Query Match 100.0%; Score 81; DB 20; Length 300;

Best Local Similarity 100.0%; Pred. No. 2.9e-06;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARMPGLSRRERFLP 16

|||||

Db 283 armpglsvrrerflp 298

RESULT 12

Y77458

ID Y77458 standard; Protein: 300 AA.

XX Y77458;

DT 05-JUN-2000 (first entry)

DE Human TNF receptor-like protein, HDTEA84.

XX TNF receptor family; tumour necrosis factor; HDTEA84; HSLJD37R;

KW Rank-like protein; RANKL: immune disorder: inflammation; allergy;

KW immunosuppressant; antiarthritic; antirheumatoid; antinflammatory;

KW dermatological; antithyroid.

XX Homo sapiens.

OS WO200001817-A2.

PN 13-JAN-2000.

XX 06-JUL-1999; 99WO-US12366.

XX 06-JUL-1998; 98US-0110938.

PR 13-JUL-1998; 98US-0114466.

PR 23-JUL-1998; 98US-0093897.

PR 12-AUG-1998; 98US-0132968.

PR 18-AUG-1998; 98US-0136214.

PR 11-SEP-1998; 98US-0099999.

XX (SCHE) SCHERING CORP.

PA Bates EEM, Lebecque SJE, Murphy EE, Mattson JD, Gorman DM;

XX Hedrick JA, Wang L, Zlotnik A, Murgolo NJ, Greene JR, Johnston JA;

PI Bazan JF, Mahony D, Lees EM;

XX WPI: 2000-171015/15.

DR N-PSDB; 292404.

XX New isolated mammalian genes, used to develop products for treating

PT e.g. immune, inflammatory or allergic abnormalities, cancers or

PT degenerative conditions -

XX Claim 24; Page 157; 218pp; English.

PS The invention relates to a number of primate and/or rodent proteins, and

XX the genes which encode them. The invention encompasses human dendritic

CC cell prostaglandin transporter (DC-PGT); the TNF (tumour necrosis

CC factor) receptor family-related proteins HDTEA84, HSLJD37R and RANKL;

CC human CC chemokine HCC5; human deubiquitinating proteins Dub1 and Dub

CC 12; human MD-1 and human and murine MD-2 proteins, which exhibit the

XX properties of ligands for proteins comprising a leucine-rich motif (LRR);

CC human cyclin E2; cDNAs encoding these proteins; and antibodies against
 CC these proteins. The proteins can be used for modulating the physiology or
 CC development of a cell. They can be used to mediate uptake of substrates
 CC (e.g., prostaglandin-like molecules), to modulate or mediate cellular
 CC interactions (e.g., induce or prevent trafficking, proliferation, or
 CC differentiation of cells), or are intracellular proteins which are
 CC important in various cellular processes such as the deubiquitination of
 CC proteins or cell cycle regulation. The products can be used for
 CC treating medical conditions such as immune, inflammatory or allergic
 CC disorders, or abnormal cellular proliferation, for example, cancers or
 CC degenerative conditions. They can be used to modulate immune responses in
 CC disease states e.g., autoimmune disorders, including rheumatoid
 CC arthritis, systemic lupus erythematosus, Hashimoto's autoimmune
 CC thyroiditis, as well as acute and chronic inflammatory responses in which
 CC T cell activation, expansion, and/or immunological T cell memory play an
 CC important role. Sequences Y77458-Y77461 and Y77465-Y77468 represent
 CC TNF receptor family-related proteins. Y77458 is the human protein
 CC HDTEA84, Y77459-Y77461 are human HSLJD37R proteins, Y77465 is
 CC murine Rank-like protein RANKL, and Y77466-Y77468 are human RANKL
 CC proteins.

XX Sequence 300 AA;

Query Match 100.0%; Score 81; DB 21; Length 300;

Best Local Similarity 100.0%; Pred. No. 2.9e-06;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARMPGLSRRERFLP 16

|||||

Db 283 armpglsvrrerflp 298

RESULT 13

Y42183

ID Y42183 standard; Protein: 302 AA.

XX Y42183;

DT 17-DEC-1999 (first entry)

DE Human FLINT #2 protein sequence.

XX Human: FLINT; mFLINT; OPG3; tumour necrosis factor receptor; Fas;
 KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
 KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
 KW reperfusion-associated injury; aplastic anaemia; differentiation;
 KW growth; myelodysplastic syndrome; pancytopenic condition;
 KW myocardial ischaemia.

OS Homo sapiens.

XX WO9950413-A2.

PN 07-OCT-1999.

XX 30-MAR-1999; 99WO-US06797.

XX 30-MAR-1998; 98US-0079856.

PR 20-MAY-1998; 98US-0086074.

PR 09-SEP-1998; 98US-0099643.

PR 17-DEC-1998; 98US-0112577.

PR 18-DEC-1998; 98US-0112703.

PR 18-DEC-1998; 98US-0112933.

PR 22-DEC-1998; 98US-0113407.

XX (ELIL) LILLY & CO ELI.

XX Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;

PI Hui KY, Kharitonkov A, Mizrahi J, Na S,

PI Song HY, Wang J, Wu X, Zuckerman SH;

XX WPI: 1999-591319/50.

DR N-PSDB; Z25376.
 XX
 XX Use of mature FLINT for treating acute liver failure, inflammation,
 PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
 PT and proinflammatory activity
 XX
 XX Example 2; Fig 2; 99pp; English.
 XX
 CC The present invention describes therapeutic applications of mature FLINT
 CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
 CC (mFLINT), which is a member of the tumour necrosis factor receptor
 CC superfamily, is used for treating acute liver failure, inflammation of
 CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
 CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
 CC injury or disorder such as hypercoagulation (including use with
 CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
 CC or disorder, Type I diabetes, cancer, cell damage or damage to an
 CC innocent bystander tissue that is induced by a chemotherapeutic agent or
 CC therapeutic irradiation, treating haematopoietic progenitor cells that
 CC have been exposed to therapeutic radiation or chemotherapy, aplastic
 CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
 CC also used for promoting the growth or differentiation of a haematopoietic
 CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
 CC resulting from abnormal myocardial ischaemia. The present sequence
 CC represents human FLINT.
 XX
 XX Sequence 302 AA;

Query Match 100.0%; Score 81; DB 20; Length 302;
 Best Local Similarity 100.0%; Pred. No. 2.9e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARMPGLERSVRERFLP 16
 |||||

Db 285 armpglersvrerflp 300

RESULT 14

ID Y84408 standard; Protein; 758 AA.
 XX Y84408;

DT 25-JUL-2000 (first entry)

DE Amino acid sequence of part of starch branching enzyme II clone 5A1.

KW Wheat; starch branching enzyme II; SBEII; SBEII-1; starch;
 KW gelatinisation onset; transgenic plant; foodstuff; bakery product.

XX Triticum aestivum.

XX Key Location/Qualifiers

FH Key
 FT Misc-difference 675 /note= "encoded by NGA"
 FT Misc-difference 680 /note= "encoded by ANA"
 FT Misc-difference 692 /note= "encoded by NTC"
 FT Misc-difference 694 /note= "encoded by ATN"
 FT Misc-difference 699 /note= "encoded by ATN"
 FT Misc-difference 709 /note= "encoded by NTA"
 FT Misc-difference 715 /note= "encoded by GNG"
 FT Misc-difference 717 /note= "encoded by NTA"
 FT Misc-difference 722 /note= "encoded by NTA"
 FT Misc-difference 722 /note= "encoded by NTA"
 FT Misc-difference 738 /note= "encoded by NTA"

FT Misc-difference /note= "encoded by NTC"
 FT 745
 FT /note= "encoded by NAC"
 FT Misc-difference 746
 FT /note= "encoded by NAA"
 FT Misc-difference 758
 FT /note= "encoded by AA"

XX WO200015810-A1.

XX 23-MAR-2000.

XX 09-SEP-1999; 99WO-GB03011.

XX 10-SEP-1998; 98EP-0307337.

XX (PLAN-) PLANT BREEDING INT CAMBRIDGE LTD.

XX Goldsbrough A, Colliver S;

XX WPI: 2000-271446/23.

XX N-PSDB; Z99938.

XX DNA encoding wheat starch branching enzyme II isoforms, useful in
 PT altering the characteristics of a plant, especially elevated starch
 PT gelatinization onset and/or peak temperature -

XX Claim 1; Fig 10; 197pp; English.

XX The present sequence represents a wheat starch branching enzyme II
 CC (SBEII) of a novel subclass of known SBEII genes, designated SBEII-1. The
 CC SBEII-1 genes are thought to have similar functional properties to the
 CC maize SBEIIb gene. Starch branching enzymes catalyse the formation of the
 CC alpha-1,6 linkages, creating branch points in the growing starch
 CC molecule, via hydrolysis of an alpha-1,4 linkage followed by reattachment
 CC of the released alpha-1,4-glucan chain to the same or another glucosyl
 CC chain. SBEII polypeptides can be used to alter the characteristics of a
 CC plant, in particular to alter starch so that it has an elevated
 CC gelatinisation onset and/or peak temperature. Starch obtained from
 CC transgenic plants is useful in the preparation or processing a foodstuff,
 CC particularly bakery products.

XX Sequence 758 AA;

Query Match 51.9%; Score 42; DB 21; Length 758;

Best Local Similarity 56.2%; Pred. No. 34;

Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 ARMPGLERSVRERFLP 16

|||||

Db 7 arypgirfgvwemflp 22

RESULT 15

W74723

ID W74723 standard; Protein; 911 AA.

XX W74723;

XX 15-FEB-1999 (first entry)

XX Human secreted protein er418_5.

XX Secreted protein; human; er418_5.

XX Homo sapiens.

XX Key Location/Qualifiers

FH Key
 FT Misc-difference 10 /note= "encoded by YTC"
 FT Misc-difference 11 /note= "encoded by YTC"
 FT

FT Misc-difference 27 /note= "encoded by YTG"
FT Misc-difference 52 /note= "encoded by CYT"
FT Domain 760
FT /note= "a potential transmembrane domain is
FT centered around amino acid 760"
XX
XX
PN WO9846757-A2.
XX
XX 22-OCT-1998.
PD
XX
XX 14-APR-1998; 98WO-US07999.
PF
XX
XX 13-APR-1998; 98US-0059487.
PR
XX 15-APR-1997; 97US-0843374.
PR
XX (GEMY) GENETICS INST INC.
PA
XX
XX Agostino MJ, Jacobs K, LaVallie ER, McCoy JM, Merberg D;
PI Racie LA, Spaulding V, Treacy M;
PI
XX
XX WPI: 1998-568731/48.
DR
XX N-PSDB; V62749.
XX
XX
PT New polynucleotide(s) encoding secreted human proteins - are derived
PT from, e.g. human foetal brain or foetal kidney cDNA libraries,
PT potentially useful as, e.g. vaccines or thrombolytic agents
XX
XX
PS Claim 19; Page 78-81; 120pp; English.
XX
XX This is the amino acid sequence of novel human secreted protein
CC er418_5, as deduced from the open reading frame of an isolated
CC human foetal brain cDNA clone (see V62749). The sequence shows
CC some similarity to database sequences. The invention provides
CC polynucleotides (see V62746-55) from foetal brain, adult testis,
CC adult brain, adult kidney and foetal kidney (all deposited as
CC ATCC 98404), which encode human secreted proteins (see W74720-29).
CC The polynucleotides and proteins are predicted to have biological
CC activities which would make them suitable for treating, preventing
CC or ameliorating medical conditions in humans and animals, although
CC no supporting data are given. Suggested activities include
CC nutritional, immune stimulating (e.g. as vaccines) or suppressing
CC activity, haematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity,
CC haemostatic and thrombolytic activity, receptor/ligand activity,
CC antiinflammatory activity, cadherin/tumour invasion suppressor
CC activity, and tumour inhibition activity.
XX
SQ Sequence 911 AA;

Query Match 50.6%; Score 41; DB 19; Length 911;
Best Local Similarity 56.7%; Pred. NO. 62;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 RMPGLRSVRER 13
Db ||| ||| |
312 rmpslrrsfrdr 323

Search completed: January 30, 2001, 16:45:27
Job time: 587 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:50:29 ; Search time 149.64 Seconds
(without alignments)
7.260 Million cell updates/sec

Title: US-09-518-931-2_COPY_283_298
Perfect score: 81
Sequence: 1 ARMPGLSVRERFLP 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_66:
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	56.8	173	2	S46355
2	46	56.8	796	2	T36308
3	45	55.6	489	2	B72767
4	44	54.3	389	2	B69096
5	42	51.9	1339	1	S20052
6	41	50.6	306	2	H75358
7	41	50.6	388	2	T36191
8	41	50.6	497	2	B82981
9	41	50.6	600	2	A33658
10	41	50.6	807	2	S54549
11	40	49.4	286	2	C72456
12	40	49.4	317	2	T39206
13	40	49.4	351	2	F81972
14	40	49.4	351	2	H81028
15	39	48.1	202	2	S30164
16	39	48.1	350	2	A55922
17	39	48.1	365	2	T34759
18	39	48.1	694	2	A82697
19	38	46.9	188	2	H70562
20	38	46.9	200	2	S44236
21	38	46.9	220	1	Q0BE49
22	38	46.9	295	1	JN0696
23	38	46.9	306	2	T03433
24	38	46.9	349	2	T46362
25	38	46.9	355	2	H71656
26	38	46.9	358	2	F71110
27	38	46.9	608	2	G83246
28	38	46.9	780	2	T03156
29	38	46.9	839	2	A56337

30 38 46.9 1181 2 A43346
31 38 46.9 2493 2 S26372
32 38 46.9 2493 2 S72349
33 37.5 46.3 414 2 T03089
34 37 45.7 129 2 T00282
35 37 45.7 210 2 S77362
36 37 45.7 272 2 H83124
37 37 45.7 337 2 S06956
38 37 45.7 360 2 F83419
39 37 45.7 378 2 A44443
40 37 45.7 383 2 S28422
41 37 45.7 405 2 A82130
42 37 45.7 417 2 T08724
43 37 45.7 467 2 S72907
44 37 45.7 515 2 C75615
45 37 45.7 600 2 A45112

1-phosphatidylinos
nonstructural poly
nonstructural poly
conserved hypochet
hypothetical prote
hypothetical prote
probable oxidoredu
segmentation prote
probable permease
basic helix-loop-h
1.4-alpha-glucan b
transcription regu
hypothetical prote
glutamyl-tRNA redu
exopolysphatase
major paraflagella

ALIGNMENTS

RESULT 1

S46355
pupI protein - Pseudomonas putida
C:Species: Pseudomonas putida
C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 08-Oct-1999
C:Accession: S46355
R:Koster, W.; van Klompenburg, W.; Bitter, W.; Leong, J.; Welsbeek, P.
EMBO J. 13, 2805-2813, 1994
A:Title: Role for the outer membrane ferric siderophore receptor PupB in signal trans
A:Reference number: S46355; MUID:94298771
A:Accession: S46355
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-173 <KOS>
A:Cross-references: EMBL:X77918; NID:g525258; PIDN:CAA54870.1; PID:g525259

Query Match 56.8%; Score 46; DB 2; Length 173;
Best Local Similarity 66.7%; Pred. No. 1.4;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 ARMPGLSVRERFL 15
| : |||| |||| ||
Db 115 AALDGLRPVREAF 129

RESULT 2

T36308
probable integral membrane ATPase - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000
C:Accession: T36308
R:Saunders, D.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M
submitted to the EMBL Data Library, March 1999
A:Reference number: Z21604
A:Accession: T36308
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-796 <SAU>
A:Cross-references: EMBL:AL035654; PIDN:CAB38587.1; GSPDB:GN00070; SCOEDB:SCE8.09
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SCE8.09
C:Superfamily: Na+/K+-transporting ATPase alpha chain; ATPase nucleotide-binding doma
F:451-618/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match 56.8%; Score 46; DB 2; Length 796;
Best Local Similarity 64.3%; Pred. No. 7;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 MPGLSVRERFLP 16

|||||

Db 497 LPCAERPVDAREFLP 510

RESULT 3

B72767 hypothetical protein APE0125 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jun-2000

C:Accession: B72767

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A:Reference number: A72450; MUID:99310339

A:Accession: B72767

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-489 <RAW>

A:Cross-references: DBJ:AP000058; NID:g5103398; PIDN:BAA79036.1; PID:d1042812; PID:g510

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE0125

C:Superfamily: Aeropyrum pernix hypothetical protein APE0125

Query Match

Best Local Similarity 55.6%; Score 45; DB 2; Length 489;

Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 PGLERSVRERFLP 16

|||||

Db 222 PGIEDKIRDEFLLP 234

RESULT 4

B69096

corrinoid/iron-sulfur protein, small subunit - Methanobacterium thermoautotrophicum (str

C:Species: Methanobacterium thermoautotrophicum

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 18-Jun-1999

C:Accession: B69096

R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;

; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني, N.

Kl, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct

A:Reference number: A69000; MUID:98037514

A:Accession: B69096

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-389 <MTH>

A:Cross-references: GB:AE000928; GB:AE000566; NID:g2622835; PIDN:AAB86184.1; PID:g262284

A:Experimental source: strain Delta H

C:Genetics:

A:Gene: MTH1712

A:Start codon: GTG

C:Superfamily: corrinoid/iron-sulfur protein small chain

Query Match

Best Local Similarity 54.3%; Score 44; DB 2; Length 389;

Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 MPGLERSVRERF 14

|||||

Db 115 MPGLPRPIREHF 126

RESULT 5

S20052

DNA-directed DNA polymerase (EC 2.7.7.7) alpha catalytic chain - Trypanosoma brucei

C:Species: Trypanosoma brucei

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: S20052; S15729

R:Leeewater, P.A.J.; Strating, M.; Murphy, N.B.; Kooy, R.F.; van der Vliet, P.C.; Ove

Nucleic Acids Res. 19, 6441-6447, 1991

A:Title: The Trypanosoma brucei DNA polymerase alpha core subunit gene is development

A:Reference number: S20051; MUID:92093600

A:Accession: S20052

A:Molecule type: DNA

A:Residues: 1-1339 <LEE>

A:Cross-references: EMBL:X60951; NID:g10499; PIDN:CAA43287.1; PID:g10501

A>Note: 828-Arg was also found

C:Genetics:

A:Gene: pola

C:Superfamily: herpesvirus DNA-directed DNA polymerase

C:Keywords: DNA binding; DNA replication; nucleotidyltransferase; nucleus; zinc finger

Query Match

Best Local Similarity 51.9%; Score 42; DB 1; Length 1339;

Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 RMPGLERSVRERFL 15

|||||

Db 497 RMPGVHRFINERAL 510

RESULT 6

H75358

LAO/AO transport system kinase - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000

C:Accession: H75358

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J

M.; Shen, H.O.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;

S.; Smith, M.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896

A:Accession: H75358

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-306 <WHI>

A:Cross-references: GB:AE002017; GB:AE000513; NID:g6459527; PIDN:NAF11302.1; PID:g645

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR1747

A:Map position: 1

Query Match

Best Local Similarity 50.6%; Score 41; DB 2; Length 306;

Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ARMPGLERSVRE 12

|:|:|:|:|:|

Db 194 ADLPGAERTVRE 205

RESULT 7

T36191

probable myo-inositol phosphate synthase - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T36191

R:Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream,

submitted to the EMBL Data Library, March 1999

A:Reference number: Z21600

A:Accession: T36191

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-388 <SAU>

A:Cross-references: EMBL:AL035707; PIDN:CAB38887.1; GSPDB:GN00070; SCOEDB:SCE29.12c

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOEDB:SCE29.12c

Query Match 50.6%; Score 41; DB 2; Length 388;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ARMPGLSVRERFLP 16
IIIIII:IIIIII

Db 198 ARLPALDELAREQGLP 213

RESULT 8

B82981
probable aldehyde dehydrogenase PA5312 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 15-Sep-2000

C:Accession: B82981

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen

A:Reference number: A82950

A:Accession: B82981

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-497 <STO>

A:Cross-references: GB:AE004944; GB:AE004091; NID:g9951628; PIDN:AAG08697.1; GSPDB:GN001

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA5312

Query Match

Best Local Similarity 50.6%; Score 41; DB 2; Length 497;

Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LERSVRERFLP 16

IIIIII:IIIIII

Db 310 VERSIKDRFLP 320

RESULT 9

A33658

parafagellar rod proteins A and B - Trypanosoma brucei

C:Species: Trypanosoma brucei

C:Date: 30-Mar-1990 #sequence_revision 30-Mar-1990 #text_change 09-Sep-1997

C:Accession: A33658; B33658; S15654

R:Schlaeppli, K.; Deflorin, J.; Seebeck, T.

J. Cell Biol. 109, 1695-1709, 1989

A:Title: The major component of the parafagellar rod of Trypanosoma brucei is a helical

A:Reference number: A33658; MUID:90009043

A:Accession: A33658

A:Molecule type: DNA

A:Residues: 1-600 <SCH>

A:Cross-references: EMBL:X14819; NID:g10483; PID:g10484

A:Genetics: PFRA

A:Accession: B33658

A:Molecule type: DNA

A:Residues: 1-600 <SC2>

A:Cross-references: EMBL:X14819; NID:g10483; PID:g10484

A:Genetics: PFRB

C:Genetics: <PFRA>

A:Gene: PFR-A

C:Genetics: <PFRB>

A:Gene: PFR-B

C:Keywords: cytoskeleton

Query Match

Best Local Similarity 50.6%; Score 41; DB 2; Length 600;

Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 RMPGLSVRERFLP 16

IIIIII:IIIIII

Db 101 RVCGLEMSVRELYKP 115

RESULT 10

S54549

probable membrane protein YMR049c - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein YMR9796.02c

C:Species: Saccharomyces cerevisiae

C:Date: 08-Jul-1995 #sequence_revision 11-Aug-1995 #text_change 26-May-2000

C:Accession: S54549

R:Devlin, K.; Churcher, C.M.

submitted to the EMBL Data Library, May 1995

A:Reference number: S54548

A:Accession: S54549

A:Molecule type: DNA

A:Residues: 1-807 <DEV>

A:Cross-references: EMBL:Z49703; NID:g817880; PID:g817882; GSPDB:GN00013; MIPS:YMR049

C:Genetics:

A:Gene: MIPS:YMR049c

A:Map position: 13R

C:Superfamily: unassigned WD repeat proteins; WD repeat homology

C:Keywords: transmembrane protein

F:433-466/Domain: WD repeat homology <WD1>

F:501-517/Domain: transmembrane #status predicted <TMM>

Query Match

Best Local Similarity 50.6%; Score 41; DB 2; Length 807;

Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 RMPGLSVRERFLP 14

IIIIII:IIIIII

Db 376 KVPYGESIRERF 388

RESULT 11

C72456

hypothetical protein APE2295 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999

C:Accession: C72456

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Ta

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero

A:Reference number: A72450; MUID:99310339

A:Accession: C72456

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-286 <KAW>

A:Cross-references: DDBJ:AP000064; NID:g5105945; PIDN:BAAB1307.1; PID:d1045093; PID:g

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE2295

Query Match

Best Local Similarity 49.4%; Score 40; DB 2; Length 286;

Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 MPGLSVRERFLP 15

IIIIII:IIIIII

Db 1 MKGLSRVRESLL 13

RESULT 12

T39206

hypothetical protein SPAC926.08c - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C;Accession: T39206
R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Ramsperger, U.; Pohl, T.
submitted to the EMBL Data Library, September 1999
A;Reference number: Z21835
A;Accession: T39206
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-317 <WOO>
A;Cross-references: EMBL:AL110469; PIDN:CAB54156.1; GSPDB:GN00066; SPDB:SPAC936.08C
A;Experimental source: strain 972h-; cosmid c926
C;Genetics:
A;Gene: SPDB:SPAC936.08c
A;Map position: 1
A;Introns: 5/2; 128/3

```

Query Match      49.4%;   Score 40;   DB 2;   Length 317;
Best Local Similarity 66.7%;   Pred. No. 28;
Matches 8;   Conservative 2;   Mismatches 2;   Indels 0;   Gaps 0;

QY      2  RMPGLERSVRE 13
      : | | | | | |
Db      286  KMKGKRSVE 297

```

RESULT 13
F81972
probable fatty acid/phospholipid biosynthesis protein NMA0542 [imported] - Neisseria meningitidis
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 19-May-2000
C:Accession: F81972
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel, P.; Holtroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream, A.; Rutherford, S.; Whitehead, P.M.A.; Woodward, J.; Barrell, B.G.
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: AB1775; UID:20222556
A:Accession: F81972
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-351 <PAR>
A:A:Cross-references: GB:AL162753; GB:AL157959; NID:g7379120; PIDN:CAB83835.1; PID:g737928
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: plsx; NMA0542
C:Superfamily: phospholipid synthesis protein

```

Query Match      49.4%; Score 40; DB 2; Length 351;
Best Local Similarity 50.0%; Pred. No. 31;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY      3  MFGLERSVRERFLP 16
          : : : : :
DB      119  IGIERRPAIAKFLP 132

```

RESULT 14
H81028 fatty acid/phospholipid synthesis protein NWB1913 [imported] - Neisseria meningitidis (9
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-May-2000
C:Accession: H81028
Rickettsin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, C.; Fleischmann, R.D.; Dougherty, B.A.;
Iri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.;
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Title: Complete number: AB1000; MUID:20175755
A:Accession: H81028
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-351 <TET>

A; Cross-references: GB:AE002540; GB:AE002098; NID:g7227159; PIDN:AAF42243.1; PID:g7227159
A; Experimental source: serogroup B, strain MMD58
C; Genetics:
A; Gene: NMB1913
C; Superfamily: phospholipid synthesis protein

```

Query Match      49.4%;      Score 40;      DB 2;      Length 351;
Best Local Similarity 50.0%; | Pred. No. 31;
Matches 7; Conservative 3; Mismatches 4; Indels
Qy      3  MPGLERSVRERFLP 16
      : : : : :
Db      119  IPGIERPAAKFLP 132

```

RESULT 15

S30164
lexA protein - Pseudomonas putida
N:Alternate names: repressor lexA
C:Species: Pseudomonas putida
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jul-2000
C:Accession: S30164
R:Garriga, X.; Calero, S.; Barber, J.
Mol. Gen. Genet. 236, 125-134, 1992
A:Title: Nucleotide sequence analysis and comparison of the lexA genes from Salmonella
A:Reference number: S30163; MUID:93156677
A:Accession: S30164
A:Molecule type: DNA
A:Residues: 1-202 <DNA>
A:Cross-references: EMBL:X63017; NID:g45338; PIDN:CAA44749.1; PID:g3980196
C:Genetics:
A:Gene: lexA
C:Superfamily: lexA repressor
C:Keywords: DNA binding; repressor; transcription regulation

```

Query Match      48.1%;      Score 39;  DB 2;  Length 202;
Best Local Similarity 53.3%;
Pred. No. 25;
Matches 8;  Conservative 1;  Mismatches 6;  Indels 0;  Gaps 0;

QY      2  RMPGLERSVREERFLP 16
      | : | | | | | | |
Db      67  RLPGLEAKAEAGLP 81

```

Search completed: January 30, 2001, 16:50:31
Job time: 721 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 17:06:04 ; Search time 80.04 Seconds
(without alignments)
6.456 Million cell updates/sec

Title: US-09-518-931-2_COPY_283_298
Perfect score: 81
Sequence: 1 ARMPGLSRRERFLP 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	42	51.9	1339	1 DPOA_TRYBB	P27727 trypanosoma
2	41	50.6	600	1 PERL_TRYBB	P22225 trypanosoma
3	41	50.6	807	1 YMT9_YEAST	Q04660 saccharomyc
4	40	49.4	936	1 CAPP_RHOPA	Q32483 rhodopsuendo
5	39	48.1	202	1 LEXA_PSEPU	P37453 pseudomonas
6	38	46.9	220	1 YALL_EBV	P03229 Epstein-Bar
7	38	46.9	295	1 GLTC_WHEAT	P16315 triticum ae
8	38	46.9	355	1 RFL_RICPR	Q9zd21 rickettsia
9	38	46.9	400	1 ER19_HUMAN	P53602 homo sapien
10	38	46.9	610	1 EDD_PSEAE	P31961 pseudomonas
11	38	46.9	839	1 PHLD_BOVIN	P80109 bos taurus
12	38	46.9	1181	1 PIP2_HUMAN	Q00722 homo sapien
13	37	45.7	337	1 HAIR_DROME	P14003 drosophila
14	37	45.7	378	1 HAIR_DROVI	P29303 drosophila
15	37	45.7	467	1 HEM1_MYCLE	P46724 mycobacteri
16	37	45.7	660	1 DNAK_CHLPN	P27542 chlamydia p
17	37	45.7	998	1 BGAL_LACIA	Q48727 lactococcus
18	36.5	45.1	376	1 FMOD_HUMAN	Q05828 homo sapien
19	36.5	45.1	376	1 FMOD_MOUSE	P50608 mus musculus
20	36.5	45.1	376	1 FMOD_RAT	P50609 rattus norv
21	36.5	45.1	755	1 ARP5_YEAST	P53946 saccharomyc
22	36	44.4	113	1 RP5M_ACICA	P33987 acinetobact
23	36	44.4	190	1 KCV_SULAC	O05982 sulfolobus
24	36	44.4	281	1 TREL_MOUSE	P15806 mus musculus
25	36	44.4	284	1 IFPL_MOUSE	P52946 mus musculus
26	36	44.4	329	1 YIID_ECOLI	P32148 escherichia
27	36	44.4	406	1 RENI_HUMAN	P00797 homo sapien
28	36	44.4	579	1 UVRC_METTH	Q26541 methanobact
29	36	44.4	638	1 PANL_RAT	P21676 rattus norv
30	36	44.4	663	1 OFRA_STRRM	Q55002 streptomyce
31	36	44.4	813	1 FPVA_PSEAE	P48632 pseudomonas
32	36	44.4	1122	1 TERT_MOUSE	O70372 mus musculus
33	36	44.4	1134	1 CHAO_DROME	P12024 drosophila

34	36	44.4	2035	1 HMP2_YEREN	P48633 versinia en
35	35.5	43.8	482	1 HOXA_ALCEU	P29267 alcaligenes
36	35	43.2	75	1 SGS8_DROME	P02842 drosophila
37	35	43.2	158	1 YNI9_YEAST	P33936 saccharomyc
38	35	43.2	166	1 Y022_GVXN	P89263 xestia c-ni
39	35	43.2	247	1 CPCI_CRYPA	P87090 cryphonectr
40	35	43.2	252	1 YE91_MYCTU	P17172 mycobacteri
41	35	43.2	307	1 GLTB_WHEAT	P10386 triticum ae
42	35	43.2	316	1 Y35K_HALSG	P17565 halobacteri
43	35	43.2	326	1 YPLE_CAUCR	P37895 caulobacter
44	35	43.2	358	1 Y993_METUA	Q58400 methanococc
45	35	43.2	359	1 YFDA_CORGL	Q99340 corynebacte

ALIGNMENTS

RESULT 1
DPOA_TRYBB
AC P27727; STANDARD; PRT; 1339 AA.
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE DNA POLYMERASE ALPHA CATALYTIC SUBUNIT (EC 2.7.7.7) (DNA POLYMERASE I).
DE Trypanosoma brucei brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S427;
RX MEDLINE=92093600; PubMed=1754381;
RA Leagwater P.A.J., Strating M.S., Murphy N.B., Kooy R.F.,
van der Vliet P.C., Overduive J.P.;
RT "The Trypanosoma brucei DNA polymerase alpha core subunit gene is
developmentally regulated and linked to a constitutively expressed
open reading frame.";
RL Nucleic Acids Res. 19:6441-6447(1991).
RC -!- FUNCTION: POLYMERASE ALPHA IN A COMPLEX WITH DNA PRIMASE IS A
CC REPLICATIVE POLYMERASE.
CC -!- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE -
CC N PYROPHOSPHATE + DNA(N).
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- MISCELLANEOUS: IN EUKARYOTES THERE ARE FIVE DNA POLYMERASES:
CC ALPHA, BETA, GAMMA, DELTA, AND EPSILON WHICH ARE RESPONSIBLE FOR
CC DIFFERENT REACTIONS OF DNA SYNTHESIS.
CC -!- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X60951; CAA43287.1; -
DR PIR: S20052; S20052.
DR INTERPRO: IPR002064; -
DR PFAM: PF00136; DNA_pol_B; 1.
DR PRINTS: PR00106; DNAPOLB.
DR PROSITE: PS00116; DNA_POLYMERASE_B; 1.
KW Transferase; DNA-directed DNA polymerase; DNA replication;
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 1339 AA; 151611 MW; 9FF159412F2B7FBA CRC64;

Query Match 51.9%; Score 42; DB 1; Length 1339;
Best Local Similarity 57.1%; Pred. No. 32;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Qy 2 RMPGLSRRERFL 15
||||:|:|:|

Db 497 RMPGVHRFINERAL 510

RESULT 2

PFRL_TRYBB STANDARD; PRT: 600 AA.

AC P22225;

DT 01-AUG-1991 (Rel. 19, Created)

DT 01-AUG-1991 (Rel. 19, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE 69 KDA PARAFAGELLAR ROD PROTEIN (69 KDA PFR PROTEIN) (PFR-A/PFR-B).

GN PFR AND PFRB.

OS Trypanosoma brucei brucei.

OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-STIB 366;

RX MEDLINE-50009043; PubMed-2793936;

RA Schieppelli K., Deflorin J., Seebeck T.;

RT "The major component of the paraflagellar rod of Trypanosoma brucei is a helical protein that is encoded by two identical, tandemly linked genes."

RL J. Cell Biol. 109:1695-1709(1989).

RN [2]

RP SEQUENCE OF 1-8 FROM N.A.

RX MEDLINE-84282716; PubMed-6088073;

RA Parsons M., Nelson R.G., Watkins K.P., Agabian N.;

RT "Trypanosome mRNAs share a common 5' spliced leader sequence."

RL Cell 38:309-316(1984).

CC -1- FUNCTION: MAJOR COMPONENT OF THE PARAFAGELLAR ROD (PFR). THE PFR IS A HIGHLY ORDERED LATTICES OF FIBROUS PROTEINS THAT ARE LOCATED INSIDE THE FLAGELLUM AND ASSUME A FIXED ORIENTATION WITH RESPECT TO THE MICROTUBULAR AXONEME.

CC -1- SUBUNIT: HETERODIMER OF A 69 KDA AND A 73 KDA PROTEIN.

CC -1- MISCELLANEOUS: CODED BY TWO TANDEMLY REPEATED GENES WITH IDENTICAL ORF: PFR AND PFRB.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

DR EMBL: X14819; CAA32930.1; -

DR EMBL: X14819; CAA32931.1; -

DR EMBL: K02195; AAA30212.1; -

DR PIR: A33658; A33658.

KW Flagella; Multigene family; Cytoskeleton; Calmodulin-binding.

FT DOMAIN 335 355 CALMODULIN-BINDING (POTENTIAL).

FT CONFLICT 7 E -> G (IN REF. 2).

SEQUENCE 600 AA; 69597 MW; CA32D77A1D7E08E1 CRC64;

Query Match 50.6%; Score 41; DB 1; Length 600;

Best Local Similarity 60.0%; Pred. NO. 20;

Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 RMPGLERSVRERFLP 16

1: ||| |||| : |

Db 101 RVCGLEMSVRELYKP 115

RESULT 3

YMT9_YEAST

ID YMT9_YEAST STANDARD; PRT: 807 AA.

AC Q04660;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE HYPOTHETICAL 91.7 KDA TRP-ASP REPEATS CONTAINING PROTEIN IN NOPL16-FAR3 INTERGENIC REGION.

DE

GN YMR049C OR YMR796.02C.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;

CC Saccharomycetaceae; Saccharomycetes.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-S288C / AB972;

RA Devlin K., Churcher C.M., Barrell B.G., Rajandream M.A.;

RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 2 WD REPEATS (TRP-ASP DOMAINS).

CC -1- SIMILARITY: STRONG, TO HUMAN KIAA0124.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

DR EMBL: Z49703; CAA89759.1; -

DR SGD: S0004652; YMR049C.

DR INTERPRO: IPR001680; -

DR PFAM: PF00400; WD40; 4.

DR PRINTS: PR00320; GPROTEINBRPT.

DR PROSITE: PS00678; WD_REPEATS; 1.

KW Hypothetical protein; Repeat; WD repeat.

SEQUENCE 807 AA; 91704 MW; 6542DFE4670341CA CRC64;

Query Match 50.6%; Score 41; DB 1; Length 807;

Best Local Similarity 53.8%; Pred. NO. 28;

Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 RMPGLERSVRERF 14

1: ||| | ||||

Db 376 KVPGEYSIRERF 388

RESULT 4

CAPP_RHOPA STANDARD; PRT: 936 AA.

AC Q32483;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE PHOSPHOENOLPYRUVATE CARBOXYLASE (EC 4.1.1.31) (PEPCASE) (PEPC).

GN PPC.

OS Rhodospseudomonas palustris.

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Bradyrhizobium group; Rhodospseudomonas.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=7;

RX MEDLINE-97386437; PubMed-9244286;

RA Inui M., Dumay V., Zahn K., Yamagata H., Yukawa H.;

RT "Structural and functional analysis of the phosphoenolpyruvate carboxylase gene from the purple nonsulfur bacterium Rhodospseudomonas palustris No. 7."

RL J. Bacteriol. 179:4942-4945(1997).

CC -1- FUNCTION: TO FORM OXALOACETATE, A FOUR-CARBON DICARBOXYLIC ACID SOURCE FOR THE TRICARBOXYLIC ACID CYCLE.

CC -1- CATALYTIC ACTIVITY: ORTHOPHOSPHATE + OXALOACETATE -> H(2)O + PHOSPHOENOLPYRUVATE + CO(2).

CC -1- PATHWAY: TRICARBOXYLIC ACID CYCLE.

CC -1- SIMILARITY: BELONGS TO THE PEPCASE FAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

```
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D89668; BAA21835.1; -.
DR INTERPRO: IPR001449; -.
DR PFAM: PF00311; PEPCase; 2.
DR PRINTS: PR00150; PEPCARXELASE.
DR PROSITE: PS00393; PEPCASE_2; 1.
DR PROSITE: PS00781; PEPCASE_1; 1.
KW Lyase; Carbon dioxide fixation; Tricarboxylic acid cycle.
FT ACT_SITE 164 164 BY SIMILARITY.
FT ACT_SITE 598 598 BY SIMILARITY.
FT ACT_SITE 936 AA; 104396 MW; 5C2CAB34AD85658D CRC64;
SQ SEQUENCE 936 AA; 104396 MW; 5C2CAB34AD85658D CRC64;

Query Match 49.4%; Score 40; DB 1; Length 936;
Best Local Similarity 72.7%; Pred. No. 48;
Matches 8; Conservative 1; Mismatches 2; Indels 2; Gaps 0;

QY 4 PGLERSVRERF 14
   I I I I I I I I
Db 881 PLLERSIRHRF 891

RESULT 5
LEXA_PSEPU
ID LEXA_PSEPU STANDARD; PRT; 202 AA.
AC P37453;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE LEXA REPRESSOR (EC 3.4.21.88).
GN LEXA.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
CC [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-93156677; PubMed=1494343;
RA Garriga X., Calero S., Barbe J.;
RT "Nucleotide sequence analysis and comparison of the lexA genes from
RT Salmonella typhimurium, Erwinia carotovora, Pseudomonas aeruginosa
RT and Pseudomonas putida.";
RL Mol. Gen. Genet. 236:125-134(1992).
CC -!- FUNCTION: REPRESENTS A NUMBER OF GENES INVOLVED IN THE RESPONSE TO
CC DNA DAMAGE (SOS RESPONSE), INCLUDING RECA AND LEXA. BINDS TO A
CC 16 BP PALINDROMIC SEQUENCE. IN THE PRESENCE OF SINGLE-STRANDED
CC DNA, RECA INTERACTS WITH LEXA CAUSING AN AUTOCATALYTIC CLEAVAGE
CC WHICH DISRUPTS THE DNA-BINDING PART OF LEXA, LEADING TO
CC DEREPRESSION OF THE SOS REGULON AND EVENTUALLY DNA REPAIR
CC (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: HYDROLYSIS OF A ALA-I-GLY BOND IN REPRESSOR
CC LEXA.
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S24 ALSO KNOWN AS THE
CC UMUD/LEXA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X63017; -. NOT_ANNOTATED_CDS.
DR PIR: S30164; S30164.
DR HSP: P03033; 1LEB.
DR MEROPS: S24.001; -.
DR INTERPRO: IPR000129; -.
DR INTERPRO: IPR002865; -.
DR PFAM: PF01726; LexA_DNA_bind; 1.
DR PFAM: PF00717; Peptidase_S24; 1.
DR PRINTS: PR00726; LEXASERPTASE.
DE GLUTENIN, LOW MOLECULAR WEIGHT SUBUNIT PTUDC1 PRECURSOR.
```

```
KW Transcription regulation; Repressor; DNA damage;
KW Autocatalytic cleavage; Hydrolase; DNA replication; SOS response;
KW DNA-binding. 28 48 H-T-H MOTIF (BY SIMILARITY).
FT ACT_SITE 88 89 CLEAVAGE (AUTO-) (BY SIMILARITY).
FT ACT_SITE 123 123 INVOLVED IN AUTO-CLEAVAGE
FT ACT_SITE 160 160 INVOLVED IN AUTO-CLEAVAGE
FT ACT_SITE 202 AA; 22147 MW; EF43C178C97D0679 CRC64;
SQ SEQUENCE 202 AA; 22147 MW; EF43C178C97D0679 CRC64;

Query Match 48.1%; Score 39; DB 1; Length 202;
Best Local Similarity 53.3%; Pred. No. 14;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 RMPGLERSVRERFLP 16
   I I I I I I I I
Db 67 RIPGLEAKAEAGLP 81

RESULT 6
YALI_EBV
ID YALI_EBV STANDARD; PRT; 220 AA.
AC P03229;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE HYPOTHETICAL BALF1 PROTEIN.
OS Epstein-Barr virus (strain B95-8) (Human herpesvirus 4).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-84270667; PubMed=6087149;
RA Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
RA Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Sequin C.,
RA Tuffnell P.S., Barrell B.G.;
RT "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
RL Nature 310:207-211(1984).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: V01555; CAA24810.1; -.
DR PIR: A03793; QQBE49.
DR PIR: S33059; S33059.
KW Hypothetical protein; Early protein.
SQ SEQUENCE 220 AA; 25148 MW; F0292CEA80626C96 CRC64;

Query Match 46.9%; Score 38; DB 1; Length 220;
Best Local Similarity 50.0%; Pred. No. 22;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 ARMPGLERSVRERF 14
   I I I I I I I I
Db 191 ARLGMPRLRRQF 204

RESULT 7
GLTC_WHEAT
ID GLTC_WHEAT STANDARD; PRT; 295 AA.
AC P16315;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GLUTENIN, LOW MOLECULAR WEIGHT SUBUNIT PTUDC1 PRECURSOR.
```

OS Triticum aestivum (wheat).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Triticum.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. MEXICALI;
RA Cassidy B.G., Dvorak J.;
RT "Molecular characterization of a low-molecular-weight glutenin CDNA
clone from Triticum durum".
RL Theor. Appl. Genet. 81:653-660(1991).
CC -!- FUNCTION: GLUTENINS ARE THE HIGH MOLECULAR WEIGHT SEED STORAGE
CC PROTEIN OF WHEAT ENDOSPERM. THOUGHT TO BE RESPONSIBLE FOR THE
CC VISCO-ELASTIC PROPERTY OF WHEAT DOUGH.
CC -!- SUBUNIT: DISULFIDE-BRIDGE LINKED AGGREGATES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X51759; CAA36063.1; -
DR PIR; S08683; S08683.
DR PIR; JN0696; JN0696.
DR INTERPRO: IPR001954; -
DR PRINTS: PR00208; GLIADGLUTEN.
KW Seed storage protein; Repeat; Multigene family; Signal.
FT SIGNAL 1 20
FT CHAIN 21 295 GLUTENIN, LOW MOLECULAR WEIGHT SUBUNIT
FT PT0UCDL.
SQ SEQUENCE 295 AA; 33379 MW; C236FBEAE740076E CRC64;

Query Match 46.9%; Score 38; DB 1; Length 295;
Best Local Similarity 57.1%; Pred. No. 31;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 3 MPGLRSVRERFLP 16
Db 26 IPGLRPWQEQPLP 39

RESULT 8
ID REF1_RICPR STANDARD; PRT; 355 AA.
AC Q9D21;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PEPTIDE CHAIN RELEASE FACTOR 1 (RF-1).
GN PREA OR RP529.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MADRID E;
RA MEDLINE=99039499; PubMed=9823893;
RX Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sacheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria".
RL Nature 396:133-140(1998).
CC -!- FUNCTION: PEPTIDE CHAIN RELEASE FACTOR 1 DIRECTS THE TERMINATION
CC OF TRANSLATION IN RESPONSE TO THE PEPTIDE CHAIN TERMINATION
CC CODONS UAG AND UAA (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE PROKARYOTIC AND MITOCHONDRIAL RELEASE
CC FACTORS FAMILY.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ235272; CAA14978.1; -
DR INTERPRO: IPR000352; -
DR PFAM; PF00472; RF-1; 1.
DR PROSITE; PS00745; RF_PROK_I; 1.
KW Protein biosynthesis.
SQ SEQUENCE 355 AA; 39895 MW; D8E3D4BAE83C67F5 CRC64;

Query Match 46.9%; Score 38; DB 1; Length 355;
Best Local Similarity 57.1%; Pred. No. 37;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 3 MPGLRSVRERFLP 16
Db 88 LPKLERAVRISLLP 101

RESULT 9
ID REF19_HUMAN STANDARD; PRT; 400 AA.
AC P53602;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE DIPHOSPHOMEVALONATE DECARBOXYLASE (EC 4.1.1.33) (MEVALONATE
DE PYROPHOSPHATE DECARBOXYLASE).
GN MPD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=96215173; PubMed=8626466;
RA Toth M.J., Huwylar L.;
RT "Molecular cloning and expression of the cDNAs encoding human and
RT yeast mevalonate pyrophosphate decarboxylase".
RL J. Biol. Chem. 271:7895-7898(1996).
CC -!- FUNCTION: PERFORMS THE FIRST COMMITTED STEP IN THE BIOSYNTHESIS OF
CC ISOPRENES.
CC -!- CATALYTIC ACTIVITY: ATP + (R)-5-DIPHOSPHOMEVALONATE -> ADP +
CC PHOSPHATE + ISOPENTENYL DIPHOSPHATE + CO(2).
CC -!- PATHWAY: CHOLESTEROL BIOSYNTHESIS.
CC -!- SUBUNIT: HOMODIMER.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN HEART, SKELETAL MUSCLE, LUNG,
CC LIVER, BRAIN, PANCREAS, KIDNEY, AND PLACENTA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U49260; AAC50440.1; -
DR MIM; 603236; -
KW Lyase; Sterol biosynthesis; Cholesterol biosynthesis.
SQ SEQUENCE 400 AA; 43404 MW; 3FDA741BCC4B68D8 CRC64;

Query Match 46.9%; Score 38; DB 1; Length 400;
Best Local Similarity 50.0%; Pred. No. 42;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9235553; PubMed=1644792;
 RA Park D., Jhon D.-Y., Kriz R., Knopf J., Rhee S.G.;
 RT "Cloning, sequencing, expression, and Gq-independent activation of
 RT phospholipase C-beta 2";
 RL J. Biol. Chem. 267:16048-16055(1992).
 CC -!- FUNCTION: THE PRODUCTION OF THE SECOND MESSENGER MOLECULES
 CC DIACYLGLYCEROL (DAG) AND INOSITOL 1,4,5-TRISPHOSPHATE (IP3) IS
 CC MEDIATED BY ACTIVATED PHOSPHATIDYLINOSITOL-SPECIFIC PHOSPHOLIPASE
 CC C ENZYMES.
 CC -!- CATALYTIC ACTIVITY: 1-PHOSPHATIDYL-D-MYO-INOSITOL 4,5-BISPHOSPHATE
 CC + H(2)O = D-MYO-INOSITOL 1,4,5-TRISPHOSPHATE + DIACYLGLYCEROL.
 CC -!- COFACTOR: REQUIRES CALCIUM.
 CC -!- MISCELLANEOUS: THERE ARE AT LEAST SIX FORMS OF PLC ENZYMES.
 CC -!- MISCELLANEOUS: THE RECEPTOR-MEDIATED ACTIVATION OF PLC-BETA-2 IS
 CC MOST EFFECTIVELY MEDIATED BY ONE G-PROTEIN ALPHA SUBUNIT,
 CC ALPHA-16.
 CC -!- SIMILARITY: DOMAINS X AND Y ARE CONSERVED IN DIFFERENT FORMS
 CC OF PLC AND ARE ESSENTIAL FOR CATALYTIC ACTIVITY.
 CC -!- SIMILARITY: CONTAINS 1 C2 DOMAIN.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; M95678; AAA36453.1; -;
 DR PIR; A43346; A43346.
 DR HSP; P10688; IDJZ.
 DR INTERPRO: IPR000008; -;
 DR INTERPRO: IPR000909; -;
 DR INTERPRO: IPR01192; -;
 DR INTERPRO: IPR01711; -;
 DR PFAM; PF00168; C2; 1.
 DR PFAM; PF00388; PI-PLC-X; 1.
 DR PFAM; PF00387; PI-PLC-Y; 1.
 DR PRINTS; PP00390; PHPLIPASEC.
 DR PROSITE; PS50004; C2_DOMAIN_2; 1.
 DR PROSITE; PS50007; PIPLC_X_DOMAIN; 1.
 DR PROSITE; PS50008; PIPLC_Y_DOMAIN; 1.
 KW Hydrolase; Lipid degradation; Transducer; Calcium.
 FT DOMAIN 312 463 DOMAIN X (BY SIMILARITY).
 FT DOMAIN 542 658 DOMAIN Y (BY SIMILARITY).
 FT DOMAIN 665 763 C2 DOMAIN.
 FT ACT_SITE 327 327 BY SIMILARITY.
 FT ACT_SITE 374 374 BY SIMILARITY.
 SQ SEQUENCE 1181 AA; 133679 MW; 7A6889F204D17FA4 CRC64;

Query Match 46.9%; Score 38; DB 1; Length 1181;
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ARMPGLERSVRE 12
 ||| ||| |||
 Db 1125 ARMKGLEAEVKE 1136

RESULT 13
 HAIR_DROME STANDARD; PRT; 337 AA.
 ID HAIR_DROME
 AC P14003;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE HAIRY PROTEIN.
 GN H.

OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Neopterygota; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=OREGON-R;
 RX MEDLINE=90059896; PubMed=2479541;
 RA Rushlow C.A., Hogan A., Pierchin S.M., Howe K.M., Lardelli M.,
 RA Ish-Horowicz D.;
 RT "The Drosophila hairy protein acts in both segmentation and bristle
 RT patterning and shows homology to N-myc";
 RL EMBO J. 8:3095-3103(1989).
 RN [2]
 RP WRPW MOTIF.
 RX MEDLINE=95094252; PubMed=8001118;
 RA Paroush Z., Finley R.L. Jr., Kidd T., Wainwright S.M., Ingham P.W.,
 RA Parent R., Ish-Horowicz D.;
 RT "Groucho is required for Drosophila neurogenesis, segmentation, and
 RT sex determination and interacts directly with hairy-related bHLH
 RT proteins";
 RL Cell 79:805-815(1994).
 CC -!- FUNCTION: PAIR-RULE PROTEIN THAT REGULATES EMBRYONIC SEGMENTATION
 CC AND ADULT BRISTLE PATTERNING. TRANSCRIPTIONAL REPRESSOR OF GENES
 CC THAT REQUIRE A BHLH PROTEIN FOR THEIR TRANSCRIPTION (EG. THE FUSHI
 CC TARAZU GENE).
 CC -!- SUBUNIT: TRANSCRIPTION REPRESSION REQUIRES FORMATION OF A COMPLEX
 CC WITH A CO-REPRESSOR PROTEIN (GROUCHO).
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.
 CC -!- DOMAIN: HAS A PARTICULAR TYPE OF BASIC DOMAIN (PRESENCE OF A
 CC HELIX-INTERRUPTING PROLINE) THAT BINDS TO THE N-BOX (CACNAG),
 CC RATHER THAN THE CANONICAL E-BOX (CANNTG).
 CC -!- DOMAIN: THE CARBOXYL-TERMINAL WRPW MOTIF IS A TRANSCRIPTIONAL
 CC REPRESSION DOMAIN NECESSARY FOR THE INTERACTION WITH GROUCHO. A
 CC TRANSCRIPTIONAL CO-REPRESSOR RECRUITED TO SPECIFIC TARGET DNA BY
 CC HAIRY-RELATED PROTEINS.
 CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 CC TRANSCRIPTION FACTORS. "HAIRY" SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X15904; CAA34018.1; -;
 DR EMBL; X15905; CAA34019.1; -;
 DR PIR; S06956; S06956.
 DR TRANSFAC; T00345; -;
 DR FLYBASE; FBgn0001168; h.
 DR INTERPRO: IPR001092; -;
 DR INTERPRO: IPR003015; -;
 DR PFAM; PF00010; HLH; 1.
 DR PROSITE; PS00038; HELIX_LOOP_HELIX; 1.
 KW Nuclear protein; Developmental protein; Pair-rule protein;
 KW DNA-binding; Transcription regulation; Repressor.
 FT DNA_BIND 32 44 BASIC DOMAIN.
 FT DOMAIN 45 89 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
 FT DOMAIN 149 157 GLN-RICH.
 FT DOMAIN 222 237 GLN-RICH.
 FT DOMAIN 241 250 POLY-ALA.
 FT VARIANT 292 292 S -> P.
 FT DOMAIN 334 337 WRPW MOTIF (REQUIRED FOR ACTIVITY).
 SQ SEQUENCE 337 AA; 36995 MW; 6D2ECAF7F2D56C0B CRC64;

Query Match 45.7%; Score 37; DB 1; Length 337;
 Best Local Similarity 46.7%; Pred. No. 52;
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ARMPGLERSVRE 15

Db 119 SRPFGIEPAQRRL 133

RESULT 14

HAIR_DROVI STANDARD; PRT; 378 AA.
AC P29303;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE HAIRY PROTEIN.
GN H.
OS Drosophila virilis (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-92269819; PubMed-1588951;
RA Mainwright S.M.; Ish-Horowicz D.;
RT "Point mutations in the Drosophila hairy gene demonstrate in vivo
requirements for basic, helix-loop-helix, and WRW domains.";
RL Mol. Cell. Biol. 12:2475-2483(1992).
CC -!- FUNCTION: PAIR-RULE PROTEIN THAT REGULATES EMBRYONIC SEGMENTATION
AND ADULT BRISTLE PATTERNING. TRANSCRIPTIONAL REPRESSOR OF GENES
THAT REQUIRE A BHLH PROTEIN FOR THEIR TRANSCRIPTION (EG. THE FUSHI
TARAZU GENE).
CC -!- SUBUNIT: TRANSCRIPTION REPRESSION REQUIRES FORMATION OF A COMPLEX
WITH A CO-REPRESSOR PROTEIN (GROUCHO).
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- DOMAIN: HAS A PARTICULAR TYPE OF BASIC DOMAIN (PRESENCE OF A
HELIX-INTERUPTING PROLINE) THAT BINDS TO THE N-BOX (CACNAG),
RATHER THAN THE CANONICAL E-BOX (CANNWG).
CC -!- DOMAIN: THE CARBOXYL-TERMINAL WRW MOTIF IS A TRANSCRIPTIONAL
REPRESSION DOMAIN NECESSARY FOR THE INTERACTION WITH GROUCHO, A
TRANSCRIPTIONAL CO-REPRESSOR RECRUITED TO SPECIFIC TARGET DNA BY
HAIRY-RELATED PROTEINS (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
TRANSCRIPTION FACTORS. "HAIRY" SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).

DR EMBL: M87885; AAA28602.1; -;
DR TRANSFAC: T01643; -;
DR FLYBASE: FBgn0013115; Dvir\h.
DR INTERPRO: IPR001092; -;
DR INTERPRO: IPR003015; -;
DR PFAM: PF00010; HLH; 1.
DR PROSITE: PS00038; HELIX LOOP HELIX; 1.
KW Nuclear protein; Developmental protein; Pair-rule protein;
KW DNA-binding; Transcription regulation; Repressor.
FT DOMAIN 22 28 POLY-GLN.
FT DNA_BIND 38 50 BASIC DOMAIN.
FT DOMAIN 51 95 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
FT DOMAIN 249 267 GLN-RICH.
FT DOMAIN 269 283 POLY-ALA.
FT DOMAIN 375 378 WRW MOTIF (REQUIRED FOR ACTIVITY).
SQ SEQUENCE 378 AA; 41383 MW; D3C2D656AA57FCAE CRC64;

Query Match 45.7%; Score 37; DB 1; Length 378;
Best Local Similarity 46.7%; Pred. No. 58;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ARMPGLERSVRREL 15

:|:|:|:|:|

Db 125 SRPFGLDSTORRL 139

RESULT 15

HEMI_MYCLE STANDARD; PRT; 467 AA.
AC P46724;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE GLUTAMYL-TRNA REDUCTASE (EC 1.2.1.-) (GLUTR).
GN HEMA OR B2168_C3_261.
OS Mycobacterium leprae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RA Smith D.R.; Robison K.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: GLUTAMYL-TRNA(GLU) + NADPH = GLUTAMATE-1-
SEMIALDEHYDE + NADP(+) + TRNA(GLU)
CC -!- PATHWAY: FIRST STEP IN PORPHYRIN BIOSYNTHESIS BY THE C5 PATHWAY.
CC -!- SIMILARITY: BELONGS TO THE GLUTAMYL-TRNA REDUCTASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).

DR EMBL: U00018; AAA17243.1; -;
DR INTERPRO: IPR000343; -;
DR PFAM: PF00745; GlutR; 1.
DR PROSITE: PS00747; GLUTR; 1.
KW Porphyrin biosynthesis; Oxidoreductase; NADP.
SQ SEQUENCE 467 AA; 49483 MW; F93245F6861F96E9 CRC64;

Query Match 45.7%; Score 37; DB 1; Length 467;
Best Local Similarity 63.6%; Pred. No. 73;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RMPGLERSVRE 12

|:|:|:|:|

Db 373 RLPGLSAQRE 383

Search completed: January 30, 2001, 17:06:06
Job time: 1227 sec

GenCore version 4.5
 Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:54:52 ; Search time 254.27 Seconds
 (without alignments)
 7.375 Million cell updates/sec

Title: US-09-518-931-2_COPY_283_298
 Perfect score: 81
 Sequence: 1 ARMPGLSVRERFLP 16

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database :

SPTREMBL_15:

- 1: sp_archaea:
- 2: sp_bacteria:
- 3: sp_fungi:
- 4: sp_human:
- 5: sp_invertebrate:
- 6: sp_mammal:
- 7: sp_mhc:
- 8: sp_organelle:
- 9: sp_phase:
- 10: sp_plant:
- 11: sp_rodent:
- 12: sp_virus:
- 13: sp_vertebrate:
- 14: sp_unclassified:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	81	100.0	300	4	Q95407
2	46	56.8	173	2	Q52208
3	46	56.8	796	2	Q924W5
4	45	55.6	489	1	Q9YFX5
5	44	54.3	389	1	Q27747
6	43	53.1	599	5	Q25310
7	41	50.6	306	2	Q9RTL3
8	41	50.6	366	10	Q9XGE9
9	41	50.6	388	2	Q924V8
10	41	50.6	600	5	Q26785
11	41	50.6	1006	5	Q9Y090
12	41	50.6	1006	5	Q9VUC7
13	40	49.4	137	2	Q9R6M0
14	40	49.4	286	1	Q9Y9J3
15	40	49.4	317	3	Q9UUG1
16	40	49.4	351	2	Q9JXR8
17	40	49.4	351	2	Q9JW54
18	40	49.4	468	5	Q9VVK1
19	39	48.1	341	2	Q9RL49

20	39	48.1	343	5	Q9VFM9
21	39	48.1	350	4	Q12792
22	39	48.1	350	11	Q09132
23	39	48.1	365	2	Q86574
24	39	48.1	392	10	Q9LMQ7
25	39	48.1	694	2	Q9PDR7
26	38.5	47.5	568	5	Q9Y005
27	38	46.9	69	2	Q51617
28	38	46.9	188	2	Q06368
29	38	46.9	200	2	Q54265
30	38	46.9	306	2	Q30543
31	38	46.9	308	2	Q9R6D7
32	38	46.9	349	4	Q9Y3F5
33	38	46.9	349	11	Q9Z0P5
34	38	46.9	358	1	Q58387
35	38	46.9	385	5	Q9VP96
36	38	46.9	432	2	Q9L629
37	38	46.9	608	12	Q66581
38	38	46.9	780	12	Q36411
39	38	46.9	1033	5	Q9N9N6
40	38	46.9	1096	5	Q9N919
41	38	46.9	1337	10	Q9LV73
42	38	46.9	1841	4	Q15031
43	38	46.9	2493	12	Q88789
44	37.5	46.3	414	12	Q55752
45	37	45.7	129	2	Q82903

ALIGNMENTS

RESULT 1

Q95407 ID O95407 PRELIMINARY: PRT: 300 AA.

AC O95407;

DT 01-MAY-1999 (TrEMBLrel. 10, Created)

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)

DE DECOY RECEPTOR 3 (M68) (M68C) (M68E).

GN DCR3 OR PR6.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99087326; PubMed=9872321;

RA Pitti R.M., Marsters S.A., Lawrence D.A., Roy M., Kischkel F.C., RA Dowd P., Huang A., Donahue C.J., Sherwood S.W., Baldwin D.T., RA Godowski P.J., Wood W.I., Gurney A.L., Hillan K.J., Cohen R.L., RA Goddard A.D., Botstein D., Ashkenazi A.

RT "Genomic amplification of a decoy receptor for Fas ligand in lung and RT colon cancer.";

RL Nature 396:699-703(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Blood;

RX MEDLINE=99253915; PubMed=10318773;

RA Yu K.Y., Kwon B., Ni J., Zhai Y., Ebner R., Kwon B.S.;

RT "A newly identified member of tumor necrosis factor receptor superfamily (TR6) suppresses LIGHT-mediated apoptosis.";

RL J. Biol. Chem. 274:13733-13736(1999).

RN [3]

RP SEQUENCE FROM N.A.

RX TISSUE=pancreas;

RC MEDLINE=20123600; PubMed=10655513;

RA Bai C., Connolly B., Metzker M.L., Hilliard C.A., Liu X., Sandig V., RA Soderman A., Galloway S.M., Liu Q., Austin C.P., Caskey C.T.;

RT "Overexpression of M68/DCR3 in human gastrointestinal tract tumors RT independent of gene amplification and its location in a four-gene RT cluster.";

RL Proc. Natl. Acad. Sci. U.S.A. 97:1230-1235(2000).

DR EMBL; AF104419; AAD03056.1; -.

Q9vfm9 drosophila
 Q12792 homo sapien
 Q09132 mus musculu
 Q86574 streptomyce
 Q9lmq7 arabidopsis
 Q9pdr7 xylella fas
 Q9y005 priapulus c
 Q51617 escherichia
 Q06368 mycobacteri
 Q54265 streptomyce
 Q30543 agrobacteri
 Q9r6d7 agrobacteri
 Q9y3f5 homo sapien
 Q9z0p5 mus musculu
 Q58387 pyrococcus
 Q9vp96 drosophila
 Q9l629 streptomyce
 Q66581 eastern equ
 Q36411 alcelaphine
 Q9n9n6 leishmania
 Q9n919 leishmania
 Q9lv73 arabidopsis
 Q15031 homo sapien
 Q88789 eastern equ
 Q55752 chilo iride
 Q82903 escherichia

DR EMBL: AF134240; AAD29688.1; -;
 DR EMBL: AF217796; AAF35244.1; -;
 DR EMBL: AF217793; AAF33685.1; -;
 DR EMBL: AF217794; AAF33686.1; -;
 DR HSSP: P25942; ICDF;
 DR INTERPRO: IPR000561; -;
 DR INTERPRO: IPR001368; -;
 DR PFAM: PF00020; TNFR_G6; 4;
 DR PROSITE: PS00652; TNFR_NGFR.1; UNKNOWN.1.
 DR PROSITE: PS01186; EGF.2; UNKNOWN.1.
 DR PROSITE: PS00050; TNFR_NGFR.2; 2;
 DR PRODOM: PD000771; -; 1;
 KW Receptor.
 SQ SEQUENCE 300 AA; 32679 MW; F90AEE33718449AF CRC64;

Query Match 100.0%; Score 81; DB 4; Length 300;
 Best Local Similarity 100.0%; Pred. No. 8.7e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ARMPGLSVRERFLP 16
 |||||
 Db 283 ARMPGLSVRERFLP 298
 |||||
 RESULT 2
 Q52208 PRELIMINARY; PRT; 173 AA.
 AC Q52208;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE PUPI AND PUPR GENES.
 GN PUPI.
 OS Pseudomonas putida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LOC5358;
 RX MEDLINE=94298771; PubMed=8026465;
 RA Koster M., Van Klompenburg W., Bitter W., Weisbeek P.;
 RT "Role for the outer membrane ferric siderophore receptor PupB in
 RT signal transduction across the bacterial cell envelope.";
 RL EMBO J. 13:2805-2813(1994).
 DR EMBL: X77918; CAA54870.1; -;
 DR INTERPRO: IPR000838; -;
 DR PFAM: PF00776; Sigma70_ECF; 1.
 SQ SEQUENCE 173 AA; 19474 MW; 8799E798E901B5DA CRC64;

Query Match 56.8%; Score 46; DB 2; Length 173;
 Best Local Similarity 66.7%; Pred. No. 3.8;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 ARMPGLSVRERFLP 15
 | : |||||
 Db 115 AALDGLRPVREAF 129

RESULT 3
 Q924W5 PRELIMINARY; PRT; 796 AA.
 AC Q924W5;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE PUTATIVE INTEGRAL MEMBRANE ATPASE.
 GN SCE8.09.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Saunders D., Harris D.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RX MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kieser H.M., Denapante D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL: AL035654; CAB38587.1; -;
 DR INTERPRO: IPR000695; -;
 DR INTERPRO: IPR001757; -;
 DR PFAM: PF00122; EI-E2_ATPase; 2.
 DR PRINTS: PR00119; CATAPASE.
 DR PROSITE: PR00120; HATPASE.
 DR PROSITE: PS00154; ATPASE_EL_E2; UNKNOWN.1.
 SQ SEQUENCE 796 AA; 83848 MW; 356868E8BAED6885 CRC64;

Query Match 56.8%; Score 46; DB 2; Length 796;
 Best Local Similarity 64.3%; Pred. No. 19;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 MPGLSVRERFLP 16
 :|||
 Db 497 LPGAERPVDARFLP 510

RESULT 4
 Q9YFX5 PRELIMINARY; PRT; 489 AA.
 ID Q9YFX5;
 AC Q9YFX5;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE HYPOTHETICAL 52.1 KDA PROTEIN APE0125.
 GN APE0125.
 OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;
 OC Aeropyrum.
 OX NCBI_TaxID=56636;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K1;
 RX MEDLINE=99310339; PubMed=10382966;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
 RA Jin-no K., Takahashi M., Sekine M., Anai A., Kosugi H.,
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
 RA Yamazaki J., Kishida N., Oguchi A., Aoki K., Kubota K., Nakamura Y.,
 RA Nomura N., Sako Y., Kikuchi H.;
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 RT crenarchaeon, Aeropyrum pernix K1.";
 RL DNA Res. 6:83-101(1999).
 DR EMBL: AF000058; BAA79036.1; -;
 DR INTERPRO: IPR003015; -;
 DR PROSITE: PS00038; HELIX_LOOP_HELIX; UNKNOWN.1.
 KW Hypothetical protein.
 SQ SEQUENCE 489 AA; 52139 MW; C7C43E0ABA7B2069 CRC64;

Query Match 55.6%; Score 45; DB 1; Length 489;

Best Local Similarity 53.8%; Pred. No. 17;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 4 PGLERSVRERFLP 16
|:|:| |:|:| |
Db 222 PGIEDKIRDEFLP 234

RESULT 5

ID O27747 PRELIMINARY; PRT; 389 AA.
AC O27747;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE CORRINOID/IRON-SULFUR PROTEIN, SMALL SUBUNIT.
GN MTH1712.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanobacterium.
OX NCBI_TaxID=21166;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DELTA H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., DeLoughery C., Lee H.-M., DuBois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwanli N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Nolling J., Reeve J.N.;
RA "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RT J. Bacteriol. 179:7135-7155(1997).
RL EMBL; AE000928; AAB86184.1; -.
DR EMBL; AE000928; AAB86184.1; -.
SQ SEQUENCE 389 AA; 43785 MW; AC75FEE22B024591 CRC64;

Query Match 54.3%; Score 44; DB 1; Length 389;
Best Local Similarity 66.7%; Pred. No. 19;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 MPGLERSVRERF 14
|:|:| |:|:| |
Db 115 MPGLPRPIREHF 126

RESULT 6

ID Q25310 PRELIMINARY; PRT; 599 AA.
AC Q25310;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE PARAFLAGELLAR ROD PROTEIN 2C.
GN PFR-2C.
OS Leishmania mexicana.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MNYC/BZ/62/M379;
RX MEDLINE=97047370; PubMed=8892290;
RA Moore L.L., Santrich C., Lebowitz J.H.;
RT "Stage-specific expression of the Leishmania mexicana paraflagellar
RL rod protein PFR-2.";
RT Mol. Biochem. Parasitol. 80:125-136(1996).
DR EMBL; U45884; AAB17719.1; -.
KW Flagella.
SQ SEQUENCE 599 AA; 68714 MW; FG20F160922002F0 CRC64;

Query Match 53.1%; Score 43; DB 5; Length 599;
Best Local Similarity 60.0%; Pred. No. 45;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 RMPGLERSVRERFLP 16
|:|:| |:|:| |
Db 100 RVAGLEMSVRELYKP 114

RESULT 7

ID Q9RTL3 PRELIMINARY; PRT; 306 AA.
AC Q9RTL3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE LAO/AO TRANSPORT SYSTEM KINASE.
GN DRI747.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RI;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RA "Genome Sequence of the Radioresistant Bacterium Deinococcus
RT radiodurans R1.";
RT Science 286:1571-1577(1999).
RL EMBL; AE002016; AAF11302.1; -.
DR TIGR; DRI747; -.
KW Kinase.
SQ SEQUENCE 306 AA; 32632 MW; B5E9F0EC50C90A22 CRC64;

Query Match 50.6%; Score 41; DB 2; Length 306;
Best Local Similarity 66.7%; Pred. No. 48;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ARMPGLERSVRE 12
|:|:| |:|:| |
Db 194 ADLPGAERTVRE 205

RESULT 8

ID Q9XGE9 PRELIMINARY; PRT; 366 AA.
AC Q9XGE9;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE LOW MOLECULAR WEIGHT GLUTENIN SUBUNIT PRECURSOR (FRAGMENT).
GN LMW-GS TYPE-1.
OS Triticum turgidum subsp. durum (durum wheat).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Triticum.
OX NCBI_TaxID=4567;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. LIRA BIOTYPE;
RA D'Ovidio R., Marchitelli C., Ercoli Cardelli L., Porceddu E.;
RT "Sequence similarity between allelic Glu-B3 genes related to quality
RT properties of durum wheat.";
RL Theor. Appl. Genet. 98:455-461(1999).
DR EMBL; Y18159; CAB41921.1; -.
DR INTERPRO; IPR000480; -.
DR INTERPRO; IPR001376; -.

OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Randell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berhan B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Boutck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laso P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.H., Zhong W., Zhou X., Zhu X., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*."
RL Science 287:2185-2195(2000).
DR EMBL; AF003536; AAF49760.1; -.
DR FLYBASE; FBgn0036409; CG6760.
DR INTERPRO; IPR001939; -.
DR PFAM; PF00004; AAA; 2.
DR PROSITE; PS00674; AAA; 1.
SQ SEQUENCE 1006 AA; 113869 MW; 4EB80F08C90394C6 CRC64;

Query Match 50.6%; Score 41; DB 5; Length 1006;
Best Local Similarity 60.0%; Pred. No. 1.7e+02;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 ARMPGLSVRERFL 15
II: I III: I I
Db 617 ARLPNERADREIL 631

RESULT 13
QYR6M0
ID QYR6M0 PRELIMINARY; PRT; 137 AA.
AC QYR6M0
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE TIORF39 PROTEIN.
GN TIORF39.

OS *Agrobacterium tumefaciens* (Ti plasmid pTiT37).
OC Plasmid pTi-SAKURA.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Agrobacterium.
OX NCBI_TaxID=362;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF301001;
RA Suzuki K., Hattori Y., Uraji M., Ohta N., Iwata K., Murata K.,
RA Kato A., Yoshida K.;
RT "Complete nucleotide sequence of a plant tumor-inducing Ti plasmid."
RL Gene 242:331-336(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF301001;
RA Suzuki K., Ohta N., Hattori Y., Uraji M., Kato A., Yoshida K.;
RT "Novel structural difference between nopaline- and octopine- type trbJ
gene: construction of genetic and physical map and sequencing of
trbJ/trai and rep gene clusters of a new Ti plasmid pTi-SAKURA."
RL Biochim. Biophys. Acta 1396:1-7(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF301001;
RA Hattori Y., Suzuki K., Ohta N., Uraji M., Kato A., Yoshida K.;
RT "Genome structure of pTi-SAKURA (I): Strategy for DNA sequencing of a
Japanese cherry-Ti plasmid."
RL Nucleic Acids Symp. Ser. 37:159-160(1998).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF301001;
RA Ohta N., Suzuki K., Hattori Y., Uraji M., Kato A., Yoshida K.;
RT "Genome structure of pTi-SAKURA (III): Characteristics of T-DNA."
RL Nucleic Acids Symp. Ser. 39:185-186(1998).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF301001;
RA Uraji M., Suzuki K., Ohta N., Hattori Y., Kato A., Yoshida K.;
RT "Genome structure of pTi-SAKURA (IV): Characteristics of tra region."
RL Nucleic Acids Symp. Ser. 39:187-188(1998).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF301001;
RA Hattori Y., Suzuki K., Ohta N., Uraji M., Kato A., Yoshida K.;
RT "Genome structure of pTi-SAKURA (V): Complete nucleotide sequence of
plasmid pTi-SAKURA's vir region in *Agrobacterium tumefaciens*."
RL Nucleic Acids Symp. Ser. 39:265-266(1998).
DR EMBL; AB016260; BAA87664.1; -.
KW Plasmid.
SQ SEQUENCE 137 AA; 15250 MW; F683EC06D0FA443A CRC64;

Query Match 49.4%; Score 40; DB 2; Length 137;
Best Local Similarity 46.7%; Pred. No. 30;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 RMPGLSVRERFLP 16
I: I I I I I I I I I I
Db 60 KAPSLERAIPARWLP 74

RESULT 14
QY9J3
ID QY9J3 PRELIMINARY; PRT; 286 AA.
AC QY9J3
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE HYPOTHETICAL 32.5 KDA PROTEIN APE2295.
GN APE2295.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;
OC Aeropyrum.
OX NCBI_TaxID=56636;

```
RN
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kwarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K., Kubota K., Nakamura Y.,
RA Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL: AP000064; BAA81307.1; -.
KW Hypothetical protein.
SQ SEQUENCE 286 AA; 32481 MW; A73BDBA7BED0E406 CRC64;

Query Match 49.4%; Score 40; DB 1; Length 286;
Best Local Similarity 69.2%; Pred. No. 66;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 MPGLERSVREREL 15
Db 1 MKGLSRSVRESLL 13

RESULT 15
Q9UUG1 PRELIMINARY; PRT; 317 AA.
AC Q9UUG1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE HYPOTHETICAL 36.1 KDA PROTEIN.
GN SPAC926.08C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;
OC Schizosaccharomycetaceae; Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972H-;
RA Wood V., Rajandream M.A., Barrell B.G., Ramsperger U., Pohl T.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL110469; CAB54156.1; -.
KW Hypothetical protein.
SQ SEQUENCE 317 AA; 36071 MW; 0438CA22A431DCF2 CRC64;

Query Match 49.4%; Score 40; DB 3; Length 317;
Best Local Similarity 66.7%; Pred. No. 73;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 RMPLERSVRER 13
Db 286 KMGLKRSVEER 297
```

Search completed: January 30, 2001, 16:54:55
Job time: 911 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:47:45 ; Search time 132.77 seconds
(without alignments)
2.164 Million cell updates/sec

Title: US-09-518-931-2_COPY_283_298

Perfect score: 81

Sequence: 1 ARMPGLSVRERFLP 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_AA.*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6_COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	81	100.0	300	2	US-08-794-796-2
2	39	48.1	350	1	US-08-184-252A-2
3	39	48.1	350	2	US-09-123-851-3
4	39	48.1	350	2	US-08-728-520-3
5	39	48.1	350	4	PCT-US95-00601-2
6	38	46.9	343	2	US-09-123-851-1
7	38	46.9	343	2	US-08-728-520-1
8	38	46.9	400	2	US-08-733-825-2
9	37	45.7	369	2	US-08-991-300-2
10	36.5	45.1	376	1	US-08-303-238-1
11	36	44.4	16	3	US-09-417-305-2
12	36	44.4	43	3	US-09-417-305-1
13	36	44.4	224	1	US-08-173-510B-87
14	36	44.4	224	1	US-08-458-218-85
15	36	44.4	224	2	US-08-450-497-87
16	36	44.4	274	1	US-08-173-510B-83
17	36	44.4	274	1	US-08-458-218-81
18	36	44.4	274	1	US-08-450-497-83
19	36	44.4	284	2	US-08-320-148B-2
20	36	44.4	284	3	US-08-589-028-6
21	36	44.4	284	3	US-08-784-582-6
22	35	43.2	15	2	US-08-726-306A-60
23	35	43.2	45	1	US-08-173-510B-15
24	35	43.2	45	1	US-08-458-218-15
25	35	43.2	45	2	US-08-450-497-15
26	35	43.2	382	3	US-08-582-740-70
27	35	43.2	401	3	US-08-289-222E-3
28	35	43.2	406	3	US-08-582-740-68

29	35	43.2	430	2	US-08-945-848-8	Sequence 8, Appli
30	35	43.2	501	2	US-08-288-508C-2	Sequence 2, Appli
31	35	43.2	570	2	US-08-967-364-1	Sequence 1, Appli
32	35	43.2	570	2	US-08-967-364-7	Sequence 7, Appli
33	35	43.2	570	3	US-09-368-408-1	Sequence 1, Appli
34	35	43.2	570	3	US-09-368-408-7	Sequence 7, Appli
35	35	43.2	4544	1	US-08-469-486-52	Sequence 52, Appli
36	35	43.2	4544	2	US-08-469-658-52	Sequence 52, Appli
37	34	42.0	117	1	US-08-249-013-6	Sequence 6, Appli
38	34	42.0	117	2	US-08-886-863-6	Sequence 6, Appli
39	34	42.0	117	4	PCT-US95-06764-6	Sequence 6, Appli
40	34	42.0	299	4	PCT-US91-00899-6	Sequence 6, Appli
41	34	42.0	340	2	US-08-794-546-5	Sequence 5, Appli
42	34	42.0	342	3	US-08-785-928-1	Sequence 1, Appli
43	34	42.0	342	3	US-08-728-603-17	Sequence 17, Appli
44	34	42.0	346	2	US-08-702-153-2	Sequence 2, Appli
45	34	42.0	346	2	US-08-702-153-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-08-794-796-2
; Sequence 2, Application US/08794796
; Patent No. 5885800
; GENERAL INFORMATION:
; APPLICANT: Emery, John
; APPLICANT: Tan, KB
; APPLICANT: Truneh, Alem
; APPLICANT: Young, Peter
; TITLE OF INVENTION: Tumor Necrosis Related Receptor,
; TITLE OF INVENTION: TR4
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/794,796
; FILING DATE: 04-FEB-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: GH50000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-4026
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 300 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-794-796-2

Query Match 100.0%; Score 81; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 9.7e-07;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ARMPGLSVRERFLP 16
| | | | | | | | | | | | | | | |
Db 283 ARMPGLSVRERFLP 298

RESULT 2

US-08-184-252A-2
; Sequence 2, Application US/08184252A
; Patent No. 5573935
; GENERAL INFORMATION:
; APPLICANT: Beeler, John F.
; APPLICANT: LaRoche, William
; APPLICANT: Aatonson, Stuart A.
; TITLE OF INVENTION: PROTEIN TYROSINE KINASE A6
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobb, Martens Olson & Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/184,252A
; FILING DATE: 18-JAN-1994

CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned A.

REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH084.001A

TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 350 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-184-252A-2

Query Match 48.1%; Score 39; DB.1; Length 350;
Best Local Similarity 61.5%; Pred. No. 23;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 MPGLSVRERFLP 15
| | | | | | | | | | | | | | | |
Db 259 MPGYTCSIRERML 271

RESULT 3

US-09-123-851-3
; Sequence 3, Application US/09123851
; Patent No. 5958405
; GENERAL INFORMATION:

APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: A NOVEL HUMAN PROTEIN TYROSINE KINASE
; NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/123.851
; FILING DATE:

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/728,520
; FILING DATE:

ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0136 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 350 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 451482

US-09-123-851-3

Query Match 48.1%; Score 39; DB 2; Length 350;
Best Local Similarity 61.5%; Pred. No. 23;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 MPGLSVRERFLP 15
| | | | | | | | | | | | | | | |
Db 259 MPGYTCSIRERML 271

RESULT 4

US-08-728-520-3
; Sequence 3, Application US/08728520
; Patent No. 5994112
; GENERAL INFORMATION:

APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: A NOVEL HUMAN PROTEIN TYROSINE KINASE
; NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/728,520
; FILING DATE: Filed Herewith

PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:

ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0136 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 350 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 451482
US-08-728-520-3

Query Match 48.1%; Score 39; DB 2; Length 350;
Best Local Similarity 61.5%; Pred. No. 23;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 MPGLERSVRERFL 15
||| |::||| |
Db 259 MPGYTCSIRERML 271

RESULT 5
PCT-US95-00601-2
Sequence 2, Application PC/TUS9500601
GENERAL INFORMATION:
APPLICANT: United States of America Department of Health and Human
APPLICANT: Services
TITLE OF INVENTION: PROTEIN TYROSINE KINASE A6
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens Olson & Bear
STREET: 620 Newport Center Drive, Sixteenth Floor
CITY: Newport Beach
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/00601
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH084.001A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 350 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-00601-2

Query Match 48.1%; Score 39; DB 4; Length 350;
Best Local Similarity 61.5%; Pred. No. 23;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 MPGLERSVRERFL 15
||| |::||| |
Db 259 MPGYTCSIRERML 271

RESULT 6
US-09-123-851-1

Sequence 1, Application US/09123851
Patent No. 5958405
GENERAL INFORMATION:
APPLICANT: Goll, Surya K.
TITLE OF INVENTION: A NOVEL HUMAN PROTEIN TYROSINE KINASE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/123,851
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/728,520
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0136 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 343 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY:
CLONE: Consensus
US-09-123-851-1

Query Match 45.9%; Score 38; DB 2; Length 343;
Best Local Similarity 53.8%; Pred. No. 34;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 3 MPGLERSVRERFL 15
||| |::||| |
Db 259 MPGYKCSIKERML 271

RESULT 7
US-08-728-520-1
Sequence 1, Application US/08728520
Patent No. 5994112
GENERAL INFORMATION:
APPLICANT: Goll, Surya K.
TITLE OF INVENTION: A NOVEL HUMAN PROTEIN TYROSINE KINASE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: U.S.
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 1.5

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/728,520
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0136 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 343 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: Consensus
; CLONE: Consensus
; US-08-728-520-1

Query Match 46.9%; Score 38; DB 2; Length 343;
Best Local Similarity 53.8%; Pred. No. 34;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 MPGLERSVRERFL 15
||| : |||
Db 259 MPGKCSIKERML 271

RESULT 8
US-08-733-825-2
; Sequence 2, Application US/08733825
; Patent No. 5837839
; GENERAL INFORMATION:
; APPLICANT: Toch, Matthew J.
; APPLICANT: Huwiler, Leslie R.
; TITLE OF INVENTION: Coding Sequences for Mevalonate
; TITLE OF INVENTION: Pyrophosphate Decarboxylase
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5837839artis Patent and Trademark Department
; STREET: 59 Route 10
; CITY: East Hanover
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07936-1080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/733,825
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/005,652
; FILING DATE: 18-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5837839ak, Henry P.
; REGISTRATION NUMBER: 33200
; REFERENCE/DOCKET NUMBER: 4-20615/P1/CGC 1834
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 277-5110
; TELEFAX: (908) 277-4606
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 400 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-733-825-2

Query Match 46.9%; Score 38; DB 2; Length 400;
Best Local Similarity 50.0%; Pred. No. 40;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 ARMPGLERSVRERFLP 16
||| : |||
Db 235 ARMAEMARCIRERDFP 250

RESULT 9
US-08-991-300-2
; Sequence 2, Application US/08991300
; Patent No. 5973225
; GENERAL INFORMATION:
; APPLICANT: D'OVIDIO, RENATO
; APPLICANT: PORCEDDU, ENRICO
; APPLICANT: MERCHITELLI, CINZIA
; APPLICANT: CARDELLI, LUISA ERCOLI
; TITLE OF INVENTION: ISOLATION AND CHARACTERIZATION OF A GENE
; TITLE OF INVENTION: ENCODING A LOW MOLECULAR WEIGHT GLUTENIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,300
; FILING DATE: 16-DEC-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IT MI 96/A 002663
; FILING DATE: 19-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 2264-0201-0X
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 369 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-991-300-2

Query Match 45.7%; Score 37; DB 2; Length 369;
Best Local Similarity 43.8%; Pred. No. 55;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 ARMPGLERSVRERFLP 16
: : ||||| : : ||
Db 1 SHIPGLRPSQQQLP 16

```
RESULT 10
US-08-303-238-1
; Sequence 1, Application US/08303238
; Patent No. 5654270
; GENERAL INFORMATION:
; APPLICANT: RUOSLAHTI, ERKKI I.
; APPLICANT: LONGAKER, MICHAEL T.
; APPLICANT: WHITBY, DAVID J.
; APPLICANT: HARPER, JOHN R.
; APPLICANT: PIERSCHBACHER, MICHAEL D.
; APPLICANT: BORDER, WAYNE A.
; TITLE OF INVENTION: INHIBITORS OF CELL REGULATORY FACTORS
; TITLE OF INVENTION: AND METHODS FOR PREVENTING OR REDUCING SCARRING
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL AND FLORES
; STREET: 4370 LA JOLLA VILLAGE DRIVE, SUITE 700
; CITY: SAN DIEGO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/303,238
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/978,931
; FILING DATE: 17-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: KONSKI, ANTOINETTE F.
; REGISTRATION NUMBER: 34,202
; REFERENCE/DOCKET NUMBER: P-LA 9453
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-535-9001
; TELEFAX: 619-535-8949
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-303-238-1

Query Match 45.18; Score 36.5; DB 1; Length 376;
Best Local Similarity 66.7; Pred. No. 69;
Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Oy 2 RMPG-LERSVRERFL 15
Db 169 RMPGPLPSRLRLHL 183

RESULT 11
US-09-417-305-2
; Sequence 2, Application US/09417305A
; Patent No. 6159723
; GENERAL INFORMATION:
; APPLICANT: MURAKAMI, Kazuo
; APPLICANT: NAKAMURA, Yukio
; APPLICANT: SUZUKI, Fumiaki
; APPLICANT: ISHIDA, Yuichi
; TITLE OF INVENTION: RENIN-ACTIVE SUBSTANCE
; FILE REFERENCE: 99_1140A
; CURRENT APPLICATION NUMBER: US/09/417.305A
; CURRENT FILING DATE: 1999-10-13
; NUMBER OF SEQ ID NOS: 4
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: peptide
US-09-417-305-2

Query Match 44.4%; Score 36; DB 3; Length 16;
Best Local Similarity 50.0%; Pred. No. 2.5;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Oy 2 RMPGLERSVRER 13
Db 5 RMPSIRESLKER 16

RESULT 12
US-09-417-305-1
; Sequence 1, Application US/09417305A
; Patent No. 6159723
; GENERAL INFORMATION:
; APPLICANT: MURAKAMI, Kazuo
; APPLICANT: NAKAMURA, Yukio
; APPLICANT: SUZUKI, Fumiaki
; APPLICANT: ISHIDA, Yuichi
; TITLE OF INVENTION: RENIN-ACTIVE SUBSTANCE
; FILE REFERENCE: 99_1140A
; CURRENT APPLICATION NUMBER: US/09/417.305A
; CURRENT FILING DATE: 1999-10-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: peptide
US-09-417-305-1

Query Match 44.4%; Score 36; DB 3; Length 43;
Best Local Similarity 50.0%; Pred. No. 7.6;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Oy 2 RMPGLERSVRER 13
Db 15 RMPSIRESLKER 26

RESULT 13
US-08-173-5108-87
; Sequence 87, Application US/08173510B
; Patent No. 5747296
; GENERAL INFORMATION:
; APPLICANT: MATTHEW MOYLE, ET AL.
; TITLE OF INVENTION: NOVEL NEUTROPHIL INHIBITORS
; NUMBER OF SEQUENCES: 104
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; CURRENT APPLICATION NUMBER: US/09/417.305A
; CURRENT FILING DATE: 1999-10-13
; NUMBER OF SEQ ID NOS: 4
; OPERATING SYSTEM: IBM P.C. DOS 5.0
```

us-09-518-931-2_copy_283_298.ra

Wed Jan 31 13:44:18 2001

```
;
;
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/173,510B
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/151,064
; FILING DATE: 10-NOV-1993
; APPLICATION NUMBER: 08/060,433
; FILING DATE: 11-MAY-1993
; APPLICATION NUMBER: 07/996,972
; FILING DATE: 24-DEC-1992
; APPLICATION NUMBER: 07/881,721
; FILING DATE: 11-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 205/073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 87:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 224 AMINO ACIDS
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PEPTIDE
; US-08-173-510B-87

Query Match 44.4%; Score 36; DB 1; Length 224;
Best Local Similarity 53.8%; Pred. No. 47;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 MPGLSVRERFL 15
   ||| |:-|:-|
Db 27 MPGFNDSIRLQFL 39

RESULT 14
US-08-458-218-85
; Sequence 85, Application US/08458218
; Patent No. 5789178
; GENERAL INFORMATION:
; APPLICANT: MATTHEW MOYLE ET AL.
; TITLE OF INVENTION: NOVEL NEUTROPHIL INHIBITORS
; NUMBER OF SEQUENCES: 102
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,218
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/151,064
; FILING DATE: 23-DEC-1993
; APPLICATION NUMBER: 08/151,064
; FILING DATE: 10-NOV-1993
; APPLICATION NUMBER: 08/060,433
; FILING DATE: 11-MAY-1993
; APPLICATION NUMBER: 07/996,972
; FILING DATE: 24-DEC-1992
; APPLICATION NUMBER: 07/881,721
; FILING DATE: 11-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 205/073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 87:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 224 AMINO ACIDS
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PEPTIDE
; US-08-173-510B-87

Query Match 44.4%; Score 36; DB 1; Length 224;
Best Local Similarity 53.8%; Pred. No. 47;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 MPGLSVRERFL 15
   ||| |:-|:-|
Db 27 MPGFNDSIRLQFL 39

RESULT 15
US-08-450-497-87
; Sequence 87, Application US/08450497
; Patent No. 5919900
; GENERAL INFORMATION:
; APPLICANT: MATTHEW MOYLE, ET AL.
; TITLE OF INVENTION: NOVEL NEUTROPHIL INHIBITORS
; NUMBER OF SEQUENCES: 104
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,497
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/173,510
; FILING DATE: 23-DEC-1993
; APPLICATION NUMBER: 08/151,064
; FILING DATE: 10-NOV-1993
; APPLICATION NUMBER: 08/060,433
; FILING DATE: 11-MAY-1993
; APPLICATION NUMBER: 07/996,972
; FILING DATE: 24-DEC-1992
; APPLICATION NUMBER: 07/881,721
; FILING DATE: 11-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 205/073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 87:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 224 AMINO ACIDS
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PEPTIDE
; US-08-458-218-85

Query Match 44.4%; Score 36; DB 1; Length 224;
Best Local Similarity 53.8%; Pred. No. 47;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 MPGLSVRERFL 15
   ||| |:-|:-|
Db 27 MPGFNDSIRLQFL 39

RESULT 15
US-08-450-497-87
; Sequence 87, Application US/08450497
; Patent No. 5919900
; GENERAL INFORMATION:
; APPLICANT: MATTHEW MOYLE, ET AL.
; TITLE OF INVENTION: NOVEL NEUTROPHIL INHIBITORS
; NUMBER OF SEQUENCES: 104
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,497
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/173,510
; FILING DATE: 23-DEC-1993
; APPLICATION NUMBER: 08/151,064
; FILING DATE: 10-NOV-1993
; APPLICATION NUMBER: 08/060,433
; FILING DATE: 11-MAY-1993
; APPLICATION NUMBER: 07/996,972
; FILING DATE: 24-DEC-1992
; APPLICATION NUMBER: 07/881,721
; FILING DATE: 11-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 205/073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 87:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 224 AMINO ACIDS
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PEPTIDE
; US-08-458-218-85
```

;
; LENGTH: 224 AMINO ACIDS
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PEPTIDE
US-08-450-497-87

Query Match 44.4%; Score 36; DB 2; Length 224;
Best Local Similarity 53.8%; Pred. No. 47;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 3 MPGLERSVRRFL 15
||| |.:||
Db 27 MPGFNDsirLQFL 39

Search completed: January 30, 2001, 16:47:46
Job time: 622 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:45:27 ; Search time 183.67 Seconds
(without alignments)
31.649 Million cell updates/sec

Title: US-09-518-931-4

Perfect score: 968

Sequence: 1 MRALGPGLSLLCLVLALPA.....PRSGRRGRCGVAGPSLAP 170

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5 ,

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq_36:*
- 1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT:*
 - 2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT:*
 - 3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT:*
 - 4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT:*
 - 5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT:*
 - 6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT:*
 - 7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT:*
 - 8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT:*
 - 9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT:*
 - 10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT:*
 - 11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT:*
 - 12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT:*
 - 13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT:*
 - 14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT:*
 - 15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT:*
 - 16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT:*
 - 17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT:*
 - 18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT:*
 - 19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT:*
 - 20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT:*
 - 21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	968	100.0	170	19	W63623 Human tumour necro
2	814	84.1	245	20	Y28449 A human tumour nec
3	814	84.1	300	19	W66102 Amino acid sequenc
4	814	84.1	300	19	W63622 Human tumour necro
5	814	84.1	300	20	Y03099 Human lung TNF-rec
6	814	84.1	300	20	Y42182 Human FLINT #1 pro
7	814	84.1	300	20	Y17479 Mammalian tumour n
8	814	84.1	300	20	Y06817 Human DCR3 polypep
9	814	84.1	300	20	W97749 Human tumour necro
10	814	84.1	300	20	W95082 Orphan receptor (H
11	803	83.0	302	20	Y42183 Human FLINT #2 pro
12	800	82.6	300	21	Y77458 Human TNF receptor

13	671	69.3	271	20	Y42184 Human mFLINT #1 pr
14	660	68.2	273	20	Y42185 Human mFLINT #2 pr
15	337	34.8	153	20	Y22222 Human TNFR superfa
16	333	34.4	215	20	W93585 Human hAPO6 protei
17	301	31.1	359	17	R99939 Mutated OCIF, OCIF
18	299.5	30.9	349	20	W83928 Human FTHMA-070 pa
19	298	30.8	401	17	R99933 Mutated OCIF, OCIF
20	297	30.7	187	17	R99950 Mutated OCIF, OCIF
21	297	30.7	197	17	R99945 Mutated OCIF, OCIF
22	297	30.7	272	17	R99944 Mutated OCIF, OCIF
23	297	30.7	321	17	R99949 Mutated OCIF, OCIF
24	297	30.7	326	17	R99940 Mutated OCIF, OCIF
25	297	30.7	327	17	R99941 Mutated OCIF, OCIF
26	297	30.7	351	17	R99943 Mutated OCIF, OCIF
27	297	30.7	393	17	R99948 Mutated OCIF, OCIF
28	297	30.7	399	17	R99942 Mutated OCIF, OCIF
29	297	30.7	401	17	R99932 Mutated OCIF, OCIF
30	297	30.7	401	17	R99934 Mutated OCIF, OCIF
31	297	30.7	401	17	R99925 Full length osteoc
32	297	30.7	401	17	R99931 Mutated OCIF, OCIF
33	297	30.7	401	19	W53239 Human OCIF genome
34	296.5	30.6	143	17	R99946 Mutated OCIF, OCIF
35	296	30.6	187	21	Y77464 Primate protein se
36	296	30.6	390	17	R99357 Human tumour necro
37	296	30.6	395	19	W57636 Modified TRL recep
38	296	30.6	401	18	W38345 Human osteoprotege
39	296	30.6	401	19	W57635 TRL receptor prote
40	296	30.6	401	20	Y05742 Tumour necrosis fa
41	296	30.6	401	20	W95030 Tumour necrosis fa
42	296	30.6	401	20	W83926 Human FTHMA-070 pr
43	296	30.6	401	21	Y43400 Osteoprotegrin pro
44	295.5	30.5	145	17	R99930 Osteoclastogenesis
45	293.5	30.3	154	17	R99929 Osteoclastogenesis

ALIGNMENTS

RESULT 1

W63623 W63623 standard; Protein; 170 AA.

XX AC W63623;

XX DT 26-OCF-1998 (first entry)

XX DE Human tumour necrosis factor receptor-6 beta protein.

XX DE Human tumour necrosis factor receptor-6 alpha; TNFR-6 alpha; TNFR-6 beta; endothelial cells; keratinocytes; normal prostate; apoptosis; prostate tumour tissue.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Peptide 1..30

FT Protein 31..170

FT Region note= "TNFR-6 beta"

FT Region 31..166

XX FT /note= "Soluble extracellular domain"

XX PN W09830694-A2.

XX PD 16-JUL-1998.

XX PF 13-JAN-1998; 98WO-US00153.

XX PR 14-JAN-1997; 97US-0035496.

XX XX (HUMA-) HUMAN GENOME SCI INC.

XX PA Ebner R, Feng P, Gentz RL, Ni J, Ruben SM, Yu G;

XX PI Ebner R, Feng P, Gentz RL, Ni J, Ruben SM, Yu G;

DR WPI: 1998-399142/34.
 DR N-PSDB; V39086.
 XX Human tumour necrosis factor receptors 6-alpha and 6-beta - used in
 PT the diagnosis of immune system-related disorder(s)
 XX
 XX Claim 20; Fig 2A-2B; 91pp; English.
 PS
 XX The present sequence represents the human tumour necrosis factor
 CC receptor-6 beta (TNFR-6 beta) protein. The invention also provides
 CC for the TNFR-6 alpha protein (W63622). TNFR-6 alpha and TNFR-6 beta are
 CC members of the tumour necrosis factor receptor (TNFR) family. TNFRs
 CC are expressed in endothelial cells, keratinocytes, normal prostate and
 CC prostate tumour tissue. For a number of disorders of these cells,
 CC particularly of the immune system, substantially altered (whether
 CC increased or decreased) levels of TNFR-6 alpha and/or TNFR-6 beta gene
 CC expression can be detected, therefore the TNFR-6 alpha and TNFR-6 beta
 CC polypeptides, nucleic acids and antibodies are claimed to be useful in
 CC the diagnosis of such disorders. Mutations of the TNFR-6 alpha and
 CC TNFR-6 beta genes can also be detected. The TNFR polypeptides are
 CC also claimed to be useful for identifying ligands which may be useful
 CC in the treatment of apoptosis related disorders.
 XX
 XX Sequence 170 AA;
 SQ

Query Match 100.0%; Score 968; DB 19; Length 170;
 Best Local Similarity 100.0%; Pred. No. 3.8e-75;
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALEGGLSLCLVLALPALLPVPVAVRGVAETPTYPWRDAETGERLVCAQCPGTFVQR 60
 DB 1 mralepgslslclvlalpallpvpavrgvaetptypwrdaetgerlvcaqcpptfvqr 60
 QY 61 PCRDSPTTCGCPPRHYTFQWNYLRCRYCNVLCGEREEARACHATHNRACRRTGFF 120
 DB 61 pcrdspttcgpcpprhytqfwnylrcrycnvlgereearachathnracrctgff 120
 QY 121 AHAGFCLHASCPGAGVIAPGSWARGAPRSGRGRCGVAGPSLAP 170
 DB 121 ahagfclhascpvgagviapgseswargaprgsgrrcgrgvgapslap 170

RESULT 2
 Y28449
 ID Y28449 standard; Protein; 245 AA.
 XX Y28449;
 AC
 XX
 DT 29-SEP-1999 (first entry)
 XX
 DE A human tumour necrosis factor-R2-like proteins (TR2P)-1.
 XX Human tumour necrosis factor-R2-like protein; TR2P; achondroplasia;
 KW osteoporosis; developmental disorder; Cushing's syndrome;
 KW muscular dystrophy; epilepsy; hereditary neuropathy;
 KW Charcot-Marie-Tooth disease; neurofibromatosis; hypothyroidism;
 KW hydrocephalus; seizure disorder; cerebral palsy; spinal bifida;
 KW congenital glaucoma; cataract; sensorineural hearing loss;
 KW reproductive disorder; infertility; ovulatory defect; endometriosis;
 KW autoimmune disorder; ectopic pregnancy; teratogenesis; spermatogenesis;
 KW immunological disorder; AIDS; Addison's disease; allergy; bronchitis;
 KW atherosclerosis; diabetes mellitus; Chron's disease; lupus;
 KW irritable bowel syndrome; multiple sclerosis; infection;
 KW neoplastic disorder; adenocarcinoma; leukaemia; lymphoma; melanoma;
 KW myeloma; sarcoma.
 KW
 XX Homo sapiens.
 OS
 XX W09931128-A2.
 PN
 XX 24-JUN-1999.
 PD
 XX

PF 02-DEC-1998; 98WO-US25649.
 XX
 PR 16-DEC-1997; 97US-0991945.
 XX
 PA (INCY-) INCYTE PHARM INC.
 XX
 PI Au-Young J, Bandman O, Hillman JL, Kaser MR, Tang YT;
 XX
 XX WPI: 1999-457916/38.
 DR N-PSDB; X89503.
 XX
 XX New tumour necrosis factor-R2-like protein - useful in the treatment
 PT of osteogenesis, developmental, reproductive, immunological and
 XX neoplastic disorders
 XX
 XX Claim 1; Fig 1A-C; 81pp; English.
 PS
 XX The present sequence represents a human tumour necrosis factor-R2-like
 CC protein (TR2P). The protein is used to treat and prevent osteogenesis,
 CC developmental, reproductive, immunological and neoplastic disorders, and
 CC also to diagnose disorders associated with TR2 protein expression. Such
 CC disorders include osteogenesis disorders such as achondroplasia and
 CC osteoporosis, developmental disorders such as Cushing's syndrome,
 CC muscular dystrophy, and epilepsy, hereditary neuropathies such as
 CC Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism,
 CC hydrocephalus, seizure disorders such as cerebral palsy and spinal
 CC bifida, congenital glaucoma, cataract, or sensorineural hearing loss,
 CC reproductive disorders such as infertility, ovulatory defects and
 CC endometriosis, autoimmune disorders, ectopic pregnancy and teratogenesis,
 CC disruption of spermatogenesis, immunological disorders such as AIDS,
 CC Addison's disease, allergies, bronchitis, atherosclerosis, diabetes
 CC mellitus, Chron's disease, lupus and irritable bowel syndrome, multiple
 CC sclerosis, viral, fungal, helminthic, parasitic and protozoal infections,
 CC and neoplastic disorders including adenocarcinoma, leukaemia, lymphoma,
 CC melanoma, myeloma, sarcoma, and teratocarcinoma.
 XX
 XX Sequence 245 AA;
 SQ

Query Match 84.1%; Score 814; DB 20; Length 245;
 Best Local Similarity 100.0%; Pred. No. 6.2e-62;
 Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALEGGLSLCLVLALPALLPVPVAVRGVAETPTYPWRDAETGERLVCAQCPGTFVQR 60
 DB 1 mralepgslslclvlalpallpvpavrgvaetptypwrdaetgerlvcaqcpptfvqr 60
 QY 61 PCRDSPTTCGCPPRHYTFQWNYLRCRYCNVLCGEREEARACHATHNRACRRTGFF 120
 DB 61 pcrdspttcgpcpprhytqfwnylrcrycnvlgereearachathnracrctgff 120
 QY 121 AHAGFCLHASCPGAGVIAPG 142
 DB 121 ahagfclhascpvgagviapg 142

RESULT 3
 W66102
 ID W66102 standard; Protein; 300 AA.
 XX W66102;
 AC
 XX 02-DEC-1998 (first entry)
 DT
 XX Amino acid sequence of tumour necrosis related receptor (TR4).
 DE Human; tumour necrosis related receptor; TR4; agonist; antagonist;
 KW inhibition; chronic; acute; inflammation; arthritis; septicemia;
 KW autoimmune disease; transplant rejection; stroke; cancer;
 KW Alzheimer's disease.
 XX
 XX Homo sapiens.
 OS
 XX

PN EP861850-A1.
 XX 02-SEP-1998.
 PD 20-JAN-1998; 98EP-0300382.
 XX 04-FEB-1997; 97US-0794796.
 PF (SWIK) SMITHKLINE BEECHAM CORP.
 PR Emery J, Tan KB, Truneh A, Young PR;
 PA WPI: 1998-508248/44.
 PI N-PSDB: V07654.
 XX New DNA encoding tumour necrosis related receptor - used to treat
 XX and prevent e.g. inflammation, arthritis, septicemia, autoimmune
 PT diseases, transplant rejection, infection, stroke, ischaemia, ARDS,
 PT restenosis, AIDS, bone disorders and cancer
 XX Claim 1; Fig 1; 21pp; English.
 PS This is the amino acid sequence of the human tumour necrosis related
 CC receptor (TR4), used in the method of the invention. The TR4 protein
 CC or its agonist can be used to treat a subject in need of enhanced
 CC TR4 polypeptide activity. The antagonist is used to inhibit TR4
 CC polypeptide activity. The active agents can be used for the
 CC treatment and prevention of diseases such as chronic and acute
 CC inflammation, arthritis, septicemia, autoimmune diseases, transplant
 CC rejection, stroke, cancer, Alzheimer's disease.
 XX Sequence 300 AA;
 SQ

Query Match 84.1%; Score 814; DB 19; Length 300;
 Best Local Similarity 100.0%; Pred. No. 7.5e-62;
 Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALEGGLSLCLVLPALLPVPVAVRGVAETPTYPWRDAETGERLVCAQCPTGFVQR 60
 DB 1 mralegpgslclvlpallpvpavrgvaetptypwrdaetgerlvcaqcpptgfvr 60
 QY 61 PCRDSPTTCGPPRHYTQFWNYLRCRYCNVLCGEREEERACHATHNRCRCRTGFF 120
 DB 61 pcrdspttcgpprhytqfwnylercrycnvlgereeeearachathnrcrcrtgff 120
 QY 121 AHAGFCLEHASCPPGAGVIAPG 142
 DB 121 ahagfclehascpvgagviapg 142

RESULT 4
 W63622
 ID W63622 standard; Protein; 300 AA.
 AC W63622;
 XX 26-OCT-1998 (first entry)
 DT Human tumour necrosis factor receptor-6 alpha protein.
 DE Human tumour necrosis factor receptor-6 alpha; TNFR-6 beta;
 KW endothelial cells; keratinocytes; normal prostate; apoptosis;
 KW prostate tumour tissue.
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT Peptide 1..30
 FT Protein 31..300
 FT /note= "TNFR-6 alpha"
 FT Region 31..282
 FT /note= "Soluble extracellular domain"

XX WO9830694-A2.
 PN 16-JUL-1998.
 PD 13-JAN-1998; 98WO-US00153.
 PF 14-JAN-1997; 97US-0035496.
 PR (HUMA-) HUMAN GENOME SCI INC.
 PA Ebner R, Feng P, Gentz RL, Ni J, Ruben SM, Yu G;
 PI WPI: 1998-399142/34.
 DR N-PSDB: V39085.
 XX Human tumour necrosis factor receptors 6-alpha and 6-beta - used in
 PT the diagnosis of immune system-related disorder(s)
 XX Claim 20; Fig 1; 91pp; English.
 PS The present sequence represents the human tumour necrosis factor
 CC receptor-6 alpha (TNFR-6 alpha) protein. The invention also provides
 CC for the TNFR-6 beta protein (W63623). TNFR-6 alpha and TNFR-6 beta are
 CC members of the tumour necrosis factor receptor (TNFR) family. TNFRs
 CC are expressed in endothelial cells, keratinocytes, normal prostate and
 CC prostate tumour tissue. For a number of disorders of these cells,
 CC particularly of the immune system, substantially altered (whether
 CC increased or decreased) levels of TNFR-6 alpha and/or TNFR-6 beta gene
 CC expression can be detected, therefore the TNFR-6 alpha and TNFR-6 beta
 CC polypeptides, nucleic acids and antibodies are claimed to be useful in
 CC the diagnosis of such disorders. Mutations of the TNFR-6 alpha and
 CC TNFR-6 beta genes can also be detected. The TNFR polypeptides are
 CC also claimed to be useful for identifying ligands which may be useful
 CC in the treatment of apoptosis related disorders.
 XX Sequence 300 AA;
 SQ

Query Match 84.1%; Score 814; DB 19; Length 300;
 Best Local Similarity 100.0%; Pred. No. 7.5e-62;
 Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALEGGLSLCLVLPALLPVPVAVRGVAETPTYPWRDAETGERLVCAQCPTGFVQR 60
 DB 1 mralegpgslclvlpallpvpavrgvaetptypwrdaetgerlvcaqcpptgfvr 60
 QY 61 PCRDSPTTCGPPRHYTQFWNYLRCRYCNVLCGEREEERACHATHNRCRCRTGFF 120
 DB 61 pcrdspttcgpprhytqfwnylercrycnvlgereeeearachathnrcrcrtgff 120
 QY 121 AHAGFCLEHASCPPGAGVIAPG 142
 DB 121 ahagfclehascpvgagviapg 142

RESULT 5
 Y03099
 ID Y03099 standard; Protein; 300 AA.
 AC Y03099;
 XX 09-DEC-1999 (first entry)
 DT Human lung TNF-receptor protein.
 DE Tumour necrosis factor; TNF; TNF receptor; human; lung; gene therapy;
 KW detection; immunoassay; diagnosis; disease; immune system; tumour;
 KW osteogenic system; cardiovascular system; central nervous system; asthma;
 KW peripheral nervous systems; transplant incompatibility; antitumor;
 KW rheumatoid arthritis; antiasthmatic; antiarthritic.
 XX Homo sapiens.
 OS

XX Key Location/Qualifiers
 FH CDS 134..1036
 FT /*tag= a
 FT /product= "TNF-receptor"
 XX
 PN DE19809978-A1.
 XX 16-SEP-1999.
 XX 09-MAR-1998; 98DE-1009978.
 XX 09-MAR-1998; 98DE-1009978.
 XX (BADI) BASF AG.
 XX Kroeger B;
 XX WPI; 1999-519473/44.
 DR N-PSDB; 209998.
 XX
 XX New soluble member of tumor necrosis factor receptor family, useful for
 PT identification specific modulators and for treating disease e.g. tumors
 PT
 PT
 XX
 PS Claim 1; Page 8-9; 10pp; German.
 XX
 CC This invention describes a novel tumor necrosis factor (TNF) receptor
 CC (I) isolated from human lung tissue. (I) is used: (i) to raise specific
 CC antibodies (Ab); (ii) to screen for specific (ant)agonists or ligands
 CC (A), potential therapeutic agents; and (iii) therapeutically (optionally
 CC expressed from a gene therapy vector) in conditions associated with a
 CC deficit of (I). Ab are used: (a) for qualitative or quantitative
 CC detection of (I) in standard immunoassays (for diagnosis of disease, or
 CC susceptibility, or for monitoring); and (b) as therapeutic inhibitors in
 CC cases where (I) is overexpressed. Nucleic acid (II) that encodes (I) is
 CC used: (A) for recombinant production of (I); (B) also its oligonucleotide
 CC fragments, in standard hybridization and/or amplification assays; (C) as
 CC source of antisense molecules or ribozymes; and (D) to produce transgenic
 CC animals (for studying (patho)physiology of (I)). Diseases possibly
 CC associated with under- or over-expression of (I) are those of the immune,
 CC osteogenic, cardiovascular and central or peripheral nervous systems, the
 CC tumors, transplant incompatibility, asthma and rheumatoid arthritis. The
 CC products of the invention have antitumor, antiasthmatic and
 CC antiarthritic activity. This sequence represents the TNF-receptor of the
 CC invention.
 XX
 SQ Sequence 300 AA;

Query Match 84.1%; Score 814; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 7.5e-62;
 Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRALEGGLSLCLVLALPALLPVPVAVRGVAETPTTPWRDAETGERLVCAQCPTGTVQR 60
 DB 1 mraleggglslclvlalpallpvpavrgvaetptypwrdaetgerlvcaqcptgtfvr 60
 QY 61 PCRRDSPPTCGPCPPRHYYTQFWNLYLRCRYCNVLCGEREEERACHATHNRCRCRTGFF 120
 DB 61 pcrrdspptcgcpcpprhytqfwnlylrcrycnvlgereeeerachathnrcrcrtgff 120
 QY 121 AHAGFCLEHASCPPGAGVIAPG 142
 DB 121 ahagfclehascpvgagviapg 142

RESULT 6
 Y42182
 ID Y42182 standard; Protein; 300 AA.
 XX
 AC Y42182;
 XX

DT 17-DEC-1999 (first entry)
 XX Human FLINT #1 protein sequence.
 DE
 XX
 KW Human; FLINT; mFLINT; OP3; tumour necrosis factor receptor; FasL;
 KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
 KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
 KW reperfusion-associated injury; aplastic anaemia; differentiation;
 KW growth; myelodysplastic syndrome; pancytopenic condition;
 KW myocardial ischaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO9950413-A2.
 XX
 PD 07-OCT-1999.
 XX
 PF 30-MAR-1999; 99WO-US06797.
 XX
 PR 30-MAR-1998; 98US-0079856.
 PR 20-MAY-1998; 98US-0086074.
 PR 09-SEP-1998; 98US-0099643.
 PR 17-DEC-1998; 98US-0112577.
 PR 18-DEC-1998; 98US-0112703.
 PR 18-DEC-1998; 98US-0112933.
 PR 22-DEC-1998; 98US-0113407.
 XX
 XX (ELIL) LILLY & CO ELI.
 PA
 XX Bumol TF, Dou S, Glasbrook AL, Gould KE, Hale JE, Heuer JG;
 PI Hui KY, Khartonenkov A, Mirzahi J, Na S, Noblitt TW, Reidy CA;
 PI Song HY, Wang J, Wu X, Zuckerman SH;
 XX
 DR WPI; 1999-591319/50.
 DR N-PSDB; 225375.
 XX
 PT Use of mature FLINT for treating acute liver failure, inflammation,
 PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
 PT and proinflammatory activity
 XX
 PS Claim 30; Fig 1; 99pp; English.
 XX
 CC The present invention describes therapeutic applications of mature FLINT
 CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
 CC (mFLINT), which is a member of the tumour necrosis factor receptor
 CC superfamily, is used for treating acute liver failure, inflammation of
 CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
 CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
 CC injury or disorder such as hypercoagulation (including use with
 CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
 CC or disorder, Type I diabetes, cancer, cell damage or damage to an
 CC innocent bystander tissue that is induced by a chemotherapeutic agent or
 CC therapeutic irradiation, treating haematopoietic progenitor cells that
 CC have been exposed to therapeutic radiation or chemotherapy, aplastic
 CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
 CC also used for promoting the growth or differentiation of a haematopoietic
 CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
 CC resulting from abnormal myocardial ischaemia. The present sequence
 CC represents human FLINT.
 XX
 SQ Sequence 300 AA;

Query Match 84.1%; Score 814; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 7.5e-62;
 Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRALEGGLSLCLVLALPALLPVPVAVRGVAETPTTPWRDAETGERLVCAQCPTGTVQR 60
 DB 1 mraleggglslclvlalpallpvpavrgvaetptypwrdaetgerlvcaqcptgtfvr 60
 QY 61 PCRRDSPPTCGPCPPRHYYTQFWNLYLRCRYCNVLCGEREEERACHATHNRCRCRTGFF 120
 DB 61 pcrrdspptcgcpcpprhytqfwnlylrcrycnvlgereeeerachathnrcrcrtgff 120

Db 61 pcrdpspttcgpcprhrytqfwnylercrycnvlgereeeearachathnrcrcrtgff 120

QY 121 AHAGFCLHASCPCGAGVIAPG 142
|||||

Db 121 ahagfclhascppgagviapg 142

RESULT 7
Y17479
ID Y17479 standard; Protein; 300 AA.
AC Y17479;
XX
XX
DT 02-AUG-1999 (first entry)
DE Mammalian tumour necrosis factor receptor OPG-2.
XX
KW Tumour necrosis factor receptor; TNF receptor; OPG-2; Paget's disease;
KW osteopenic disorder; osteoclast activity; primary osteoporosis;
KW hyperglycaemia; osteolytic metastasis; immune response; cancer.
XX
OS Mammalia.
XX
XX WO9926977-A1.
XX
XX 03-JUN-1999.
XX
XX 24-NOV-1998; 98WO-US25065.
XX
XX 17-FEB-1998; 98US-0074896.
XX
XX 24-NOV-1997; 97US-0066446.
XX
XX (BIOJ) BIOGEN INC.
XX
XX Tschoep J;
XX
XX WPI; 1999-347693/29.
XX
XX N-PSDB; X76052.
XX
XX
PT New tumour necrosis factor family receptor OPG-2
XX
XX
PS Claim 1; Page 18; 22pp; English.
XX
XX The present sequence represents a mammalian tumour necrosis factor
XX receptor, designated OPG-2. OPG-2, is a member of the tumour necrosis
XX factor receptor family, and can be used: (i) to raise specific
XX antibodies (Ab), (ii) to treat osteopenic disorders associated with
XX excessive osteoclast activity, e.g. primary osteoporosis, Paget's
XX disease, hyperglycaemia of malignancy, or osteolytic metastases; (iii)
XX for affinity purification of cognate ligands, and (iv) to screen for
XX ligands (antagonists or agonists). Ab, or other OPG-2 blocking agents
XX such as soluble forms of the protein, are used to prevent, or reduce
XX severity of, an immune response, and for treating cancer. They can also
XX be used in diagnostic assays. The nucleic acid sequence encoding OPG-2
XX can be used as a probe to isolate related sequences from other species.
XX
SQ Sequence 300 AA;

Query Match 84.1%; Score 814; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 7.5e-62;
Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRAEGPGLSLCLVLPALLPVPVARGVAETPTYPWRDAETGERLVCAQCPGTFVQR 60
|||||

Db 1 mralegpglslclvialpallpvpavrgvaetptyprwdaetgerlvcaqcpptfvqr 60
|||||

QY 61 PCRRDSPPTTCGPPRHYTQFWNLYLRCRYCNVLCGEREEEARACHATHNRCRCRTGFF 120
|||||

Db 61 pcrdpspttcgpcprhrytqfwnylercrycnvlgereeeearachathnrcrcrtgff 120
|||||

QY 121 AHAGFCLHASCPCGAGVIAPG 142
|||||

Db 121 ahagfclhascppgagviapg 142

RESULT 8
Y06817
ID Y06817 standard; Protein; 300 AA.
XX
XX Y06817;
XX
XX 24-JUN-1999 (first entry)
XX
XX Human Dcr3 polypeptide.
XX
XX Dcr3 polypeptide; tumour necrosis factor receptor; TNFR; Fas ligand;
KW apoptosis; T cell mediated immune response; allergy; asthma; cancer;
KW rheumatoid arthritis; Crohn's disease; guest vs. host disease; human;
KW gene therapy.
XX
XX Homo sapiens.
XX
XX WO9914330-A1.
XX
XX 25-MAR-1999.
XX
XX 18-SEP-1998; 98WO-US19661.
XX
XX 30-JUL-1998; 98US-0094640.
XX
XX 18-SEP-1997; 97US-0059288.
XX
XX (GETH) GENENTECH INC.
XX
XX Ashkenazi AJ, Botstein D, Dodge KH, Goddard A, Gurney AL;
XX Kim KJ, Lawrence DA, Pitti R, Roy MA, Tumas DB;
XX Wood WI;
XX
XX WPI; 1999-244032/20.
XX
XX N-PSDB; X32744.
XX
XX Dcr3 polypeptide related to tumor necrosis factor receptor
XX
XX Claim 5; Fig 1; 88pp; English.
XX
XX This represents a human Dcr3 polypeptide, a homologue of tumour necrosis
XX factor receptor (TNFR) polypeptide. Host cells containing a vector
XX comprising the Dcr3 nucleic acid can be used for the recombinant
XX expression of the protein. Dcr3 binds to Fas ligand, so it (or its
XX chimeras) are useful for modulating apoptosis in mammalian cells, also
XX other Fas-ligand induced activities, particularly to inhibit T cell
XX mediated immune responses, e.g. in treatment of allergy, asthma,
XX rheumatoid arthritis, Crohn's disease, guest vs. host disease etc. Dcr3
XX may also be used to identify specific binding proteins, potential
XX inhibitors. Antibodies against Dcr3 are used to treat cancer,
XX specifically of the lung and colon, also in diagnosis and for affinity
XX purification of the protein. Detecting mutations in the gene for Dcr3 is
XX also used to diagnose cancer, or predisposition to it. Dcr3 nucleic acid
XX is useful as hybridization probe to detect genomic or related sequences;
XX for chromosome and gene mapping; as source of antisense sequences; for
XX expression of recombinant Dcr3 and to generate transgenic animals (for
XX ex vivo gene therapy.
XX
SQ Sequence 300 AA;

Query Match 84.1%; Score 814; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 7.5e-62;
Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRAEGPGLSLCLVLPALLPVPVARGVAETPTYPWRDAETGERLVCAQCPGTFVQR 60
|||||

Db 1 mralegpglslclvialpallpvpavrgvaetptyprwdaetgerlvcaqcpptfvqr 60
|||||

QY 61 PCRRDSPPTTCGPPRHYTQFWNLYLRCRYCNVLCGEREEEARACHATHNRCRCRTGFF 120
|||||

CC regulation and to regulate the maturation of TNF ligand-bearing
CC cells such as T- or B-cells, lymphocytes, peripheral blood
CC mononuclear cells, polymorphonuclear leukocytes, fibroblasts or
CC hematopoietic cells.
XX
SQ Sequence 300 AA;

Query Match 84.1%; Score 814; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 7.5e-62;
Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALEGPGCLSLCLVLPVAVRGVAETPTYPWRDAETGERLVCAQCPTGTVQR 60
Db 1 mralegpgslslclvlalpallpvpavrgvaetptypwrdaetgerlvcaqcpgtgvqr 60
QY 61 PCRDSPTTCGPPRHYTQFNYLERCRVCNVLCGEREEERACHATHNRCRCRTGFF 120
Db 61 pcrdspttcgpprhytqfwnylcrvcnvlcrvncvlgereearachathnrcrcrtgff 120
QY 121 AHAGFCLHASCPPGAGVIAPG 142
Db 121 ahagfclhascpvgagviapg 142

RESULT 10
W95082
ID W95082 standard; Protein: 300 AA.
XX
AC W95082;
XX
DT 20-MAY-1999 (first entry)
XX
DE Orphan receptor (HUMAN NTR-1) polypeptide.
XX
KW HUMAN NTR-1; orphan receptor; osteoprotegerin; OPG; TNFR; human;
KW tumour necrosis factor receptor; muscle disorder; bone mass; screening;
KW muscle metabolism; binding agent; cognate ligand.
XX
OS Homo sapiens.
XX
PN WO9907738-A2.
XX
PD 18-FEB-1999.
XX
PE 04-AUG-1998; 98WO-US16202.
XX
PR 06-AUG-1997; 97US-0054869.
XX
PR (PROC) PROCTER & GAMBLE CO.
PA (REGE-) REGENERON PHARM INC.
XX
PI Maslakowski PJ, Morris J, Valenzuela DM;
XX
DR WPI: 1999-167365/14.
DR N-PSDB; X22300.
XX
PT Novel orphan human receptor polypeptide and nucleic acid - useful as
XX diagnostic reagents and for treatment of muscle disorders
XX
PS Claim 7; Page 21; 23pp; English.
XX
CC This represents a HUMAN NTR-1 polypeptide, a novel orphan receptor. The
CC protein is related to osteoprotegerin (OPG) and to tumour necrosis factor
CC receptor (TNFR). Host cells transformed with a vector comprising the
CC HUMAN NTR-1 nucleic acid are used for the recombinant expression of the
CC protein. HUMAN NTR-1 proteins and antibodies immuno specific for the
CC protein are useful for diagnosis and treatment of humans and animals,
CC especially muscle disorders, as the receptor is involved in regulation of
CC bone mass and muscle metabolism. HUMAN NTR-1 receptors are also useful
CC for screening for novel binding agents, and cognate ligands, which may be
CC used to treat disorders associated with HUMAN NTR-1 imbalance.

Db 61 pcrdspttcgpprhytqfwnylcrvcnvlcrvncvlgereearachathnrcrcrtgff 120
QY 121 AHAGFCLHASCPPGAGVIAPG 142
Db 121 ahagfclhascpvgagviapg 142

RESULT 9
W97749
ID W97749 standard; Protein: 300 AA.
XX
AC W97749;
XX
DT 21-MAY-1999 (first entry)
XX
DE Human tumour necrosis factor receptor ZTNFR-5.
XX
KW ZTNFR-5; tumour necrosis factor receptor; TNFR; human;
KW cell maturation; bone cell regulation.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..23
FT Protein 24..300
FT Domain 24..194
FT Region 49..71 "cysteine-rich pseudo-repeat 1"
FT Region 72..113
FT Region 114..151
FT Region 152..194
FT Region 152..194 "cysteine-rich pseudo-repeat 1"
XX
PN WO9904001-A1.
XX
PD 28-JAN-1999.
XX
PF 21-JUL-1998; 98WO-US15072.
XX
PR 21-JUL-1997; 97US-0053203.
XX
PR (ZYMO) ZYMOGENETICS INC.
XX
PI Farrah TM;
XX
DR WPI: 1999-132245/11.
DR N-PSDB; X07226.
XX
PT Novel tumour necrosis factor receptor ZTNFR5 - useful for
XX regulating maturation of TNF-ligand bearing cells
XX
PS Claim 1; Page 84-85; 109pp; English.
XX
CC This polypeptide comprises a new, secreted tumour necrosis factor
CC receptor (see W97749), designated ZTNFR-5. Novel ZTNFR-5 encoding
CC polynucleotides and polypeptides were initially identified by
CC querying an expressed sequence tag (EST) database for sequences
CC homologous to conserved motifs within the TNF receptor family.
CC Based on this search, a contig of 16 ESTs (see X07226) was
CC constructed. ZTNFR-5 polypeptides comprise 4 cysteine-rich repeats
CC (see also W97750-55) that are homologous to other TNF receptors, in
CC particular the soluble, secreted TNF receptor osteoprotegerin.
CC ZTNFR-5 polypeptide can be prepared by recombinant methods. The
CC polypeptide, especially the extracellular domain, can be used to
CC generate a soluble variant of ZTNFR-5. The polypeptides and
CC nucleic acids can be used to screen for ligands, agonists and
CC antagonists of ZTNFR-5. The polypeptides can be used in bone cell

SQ Sequence 300 AA;

Query Match 84.1%; Score 814; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 7.5e-62;
 Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALEGPSLCLVLALPALLPVPVAVRGVAETPTYPWRDAETGERLVCAQCPGTFVQR 60
 |||||
 DB 1 mraleppglsllclvlalpalpvpavrgvaetpypwrdaetgerlvcaqcpptfvqr 60
 |||||

QY 61 PCRRDSPPTTCGPPRHYTQFWNYLRCRYCNVLCGEREEARACHATHNRCRCRTGFF 120
 |||||
 DB 61 pccrdspttcgpprhytqfwnylercrycnvlgereearachathnrcrcrtgff 120
 |||||

QY 121 AHAGFCLEHASCPPGAGVIAPG 142
 |||||
 DB 121 ahagfclehascpvgagiapg 142
 |||||

RESULT 11
 Y42183
 ID Y42183 standard; Protein; 302 AA.
 AC Y42183;
 XX
 XX
 DT 17-DEC-1999 (first entry)
 XX
 DE Human FLINT #2 protein sequence.
 XX
 KW Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
 KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
 KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
 KW reperfusion-associated injury; aplastic anaemia; differentiation;
 KW growth; myelodysplastic syndrome; pancytopenic condition;
 KW myocardial ischaemia.
 KW
 XX Homo sapiens.
 OS
 XX
 PN W099504113-A2.
 XX
 PD 07-OCT-1999.
 XX
 XX 30-MAR-1999; 99WO-US06797.
 XX
 XX 30-MAR-1998; 98US-0079856.
 PR 20-MAY-1998; 98US-0086074.
 PR 09-SEP-1998; 98US-0099643.
 PR 17-DEC-1998; 98US-0112577.
 PR 18-DEC-1998; 98US-0112703.
 PR 18-DEC-1998; 98US-0112933.
 PR 22-DEC-1998; 98US-0113407.
 XX
 XX (ELIL) LILLY & CO ELI.
 XX
 XX Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
 PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
 PI Song HY, Wang J, Wu X, Zuckerman SH;
 XX
 XX WPI; 1999-591319/50.
 DR N-PSDB; 225376.
 XX
 XX Use of mature FLINT for treating acute liver failure, inflammation,
 PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
 PT and proinflammatory activity
 XX
 XX Example 2; Fig 2; 99pp; English.
 XX
 CC The present invention describes therapeutic applications of mature FLINT
 CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
 CC (mFLINT), which is a member of the tumour necrosis factor receptor
 CC superfamily, is used for treating acute liver failure, inflammation of
 CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated

CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
 CC injury or disorder such as hypercoagulation (including use with
 CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
 CC or disorder, Type I diabetes, cancer, cell damage or damage to an
 CC innocent bystander tissue that is induced by a chemotherapeutic agent or
 CC therapeutic irradiation, treating haematopoietic progenitor cells that
 CC have been exposed to therapeutic radiation or chemotherapy, aplastic
 CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
 CC also used for promoting the growth or differentiation of a haematopoietic
 CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
 CC resulting from abnormal myocardial ischaemia. The present sequence
 CC represents human FLINT.
 XX
 SQ Sequence 302 AA;

Query Match 83.0%; Score 803; DB 20; Length 302;
 Best Local Similarity 98.6%; Pred. No. 6.5e-61;
 Matches 142; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

QY 1 MRALEGPSLCLVLALPALLPVPVAVRGVAETPTYPWRDAETGERLVCAQCPGTFVQR 60
 |||||
 DB 1 mraleppglsllclvlalpalpvpavrgvaetpypwrdaetgerlvcaqcpptfvqr 60
 |||||

QY 61 PCRRDSPPTTCGPPRHYTQFWNYLRCRYCNVLCGEREEARACHATHNRA--CRCRTG 118
 |||||
 DB 61 pccrdspttcgpprhytqfwnylercrycnvlgereearachathnrcrcrtg 120
 |||||

QY 119 FFAHAGFCLEHASCPPGAGVIAPG 142
 |||||
 DB 121 ffahagfclehascpvgagiapg 144
 |||||

RESULT 12
 Y77458
 ID Y77458 standard; Protein; 300 AA.
 XX
 AC Y77458;
 XX
 DT 05-JUN-2000 (first entry)
 XX
 DE Human TNF receptor-like protein, HDTEA84.
 XX
 KW TNF receptor family; tumour necrosis factor; HDTEA84; HSLJD37R;
 KW Rank-like protein; RANKL; immune disorder; inflammation; allergy;
 KW immunosuppressant; antiarthritic; antirheumatoid; antiinflammatory;
 KW dermatological; antithyroid.
 XX
 OS Homo sapiens.
 XX
 PN W0200001817-A2.
 XX
 PD 13-JAN-2000.
 XX
 XX 06-JUL-1999; 99WO-US12366.
 XX
 PR 06-JUL-1998; 98US-0110938.
 PR 13-JUL-1998; 98US-0114466.
 PR 23-JUL-1998; 98US-0093897.
 PR 12-AUG-1998; 98US-0132968.
 PR 18-AUG-1998; 98US-0136214.
 PR 11-SEP-1998; 98US-0099999.
 XX
 XX (SCHE) SCHERING CORP.
 PA
 XX
 XX Bates EEM, Lebecque SJE, Murphy EE, Mattson JD, Gorman DM;
 PI Hedrick JA, Wang L, Zlotnik A, Murgolo NJ, Greene JR, Johnston JA;
 PI Bazan JF, Mahony D, Lees EM,
 XX
 XX WPI; 2000-171015/15.
 DR N-PSDB; 292404.
 XX
 PT New isolated mammalian genes, used to develop products for treating

PT e.g. immune, inflammatory or allergic abnormalities, cancers or
PT degenerative conditions

XX Claim 24; Page 157; 218pp; English.

XX The invention relates to a number of primate and/or rodent proteins, and
CC the genes which encode them. The invention encompasses human dendritic
CC cell prostaglandin transporter (DC-PGT); the TNF (tumour necrosis
CC factor) receptor family-related proteins HPTA84, HSLJD37R and RANKL;
CC human CC chemokine HCC5; human deubiquitinating proteins Dub11 and Dub
CC 12; human MD-1 and human and murine MD-2 proteins, which exhibit the
CC properties of ligands for proteins comprising a leucine-rich motif (LRR);
CC human cyclin E2; cDNAs encoding these proteins; and antibodies against
CC these proteins. The proteins can be used for modulating the physiology or
CC development of a cell. They can be used to mediate uptake of substrates
CC (e.g., prostaglandin-like molecules), to modulate or mediate cellular
CC interactions (e.g., induce or prevent trafficking, proliferation, or
CC differentiation of cells), or are intracellular proteins which are
CC important in various cellular processes such as the deubiquitination of
CC proteins or cell cycle regulation. The products can be used for
CC treating medical conditions such as immune, inflammatory or allergic
CC disorders, or abnormal cellular proliferation, for example, cancers or
CC degenerative conditions. They can be used to modulate immune responses in
CC disease states e.g., autoimmune disorders, including rheumatoid
CC arthritis, systemic lupus erythematosus, Hashimoto's autoimmune
CC thyroiditis, as well as acute and chronic inflammatory responses in which
CC T cell activation, expansion, and/or immunological T cell memory play an
CC important role. Sequences Y77458-Y77461 and Y77465-Y77468 represent
CC TNF receptor family-related proteins. Y77458 is the human protein
CC HPTA84, Y77459-Y77461 are human HSLJD37R proteins, Y77465 is
CC murine Rank-like protein RANKL, and Y77466-Y77468 are human RANKL
CC proteins.

XX Sequence 300 AA;

Query Match 82.6%; Score 800; DB 21; Length 300;
Best Local Similarity 98.6%; Pred. No. 1-2e-60;
Matches 140; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRALGPGSLCLVLPALLPVPVAVGVAETPTVWRDAETGERLVCAQCPPTGVQR 60
DB 1 mralgpgslclvlpallpvpvavgvaetptywrdaetgerlvcaqcpptgvqr 60
QY 61 PCRRDSPTCGPPRHVYTFQWNYLRCRYCNVLCGEREEERACHATHNRCRCRTGFF 120
DB 61 pccrrdspmtcgprrhytqfwnylrcrycylcgereeeearachathnrcrctgff 120
QY 121 AHAGFCLHASCPPGAGVIAPG 142
DB 121 ahagfclhascppgagviapg 142

RESULT 13
Y42184
ID Y42184 standard; Protein: 271 AA.

XX Y42184;

XX 17-DEC-1999 (first entry)

XX Human mFLINT #1 protein sequence.

XX Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
KW reperfusion-associated injury; aplastic anaemia; differentiation;
KW growth; myelodysplastic syndrome; pancytopenic condition;
KW myocardial ischaemia.

OS Homo sapiens.

XX WO9950413-A2.

PN

XX 07-OCT-1999.

XX 30-MAR-1999; 99WO-US06797.

XX 30-MAR-1998; 98US-0079856.
XX 20-MAY-1998; 98US-0086074.
XX 09-SEP-1998; 98US-0099643.
XX 17-DEC-1998; 98US-0112577.
XX 18-DEC-1998; 98US-0112703.
XX 18-DEC-1998; 98US-0112933.
XX 22-DEC-1998; 98US-0113407.

XX (ELIL) LILLY & CO ELI.

XX Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
PI Song HY, Wang J, Wu X, Zuckerman SH;

XX WPI: 1999-591319/50.

XX N-PSDB; 225377.

XX Use of mature FLINT for treating acute liver failure, inflammation,
XX cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
XX and proinflammatory activity

XX Claim 31; Fig 3; 99pp; English.

XX The present invention describes therapeutic applications of mature FLINT
XX (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
XX (mFLINT), which is a member of the tumour necrosis factor receptor
XX superfamily, is used for treating acute liver failure, inflammation of
XX the liver, abnormal hepatocyte apoptosis, sepsis, an ischaemia-associated
XX with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
XX injury or disorder such as hypercoagulation (including use with
XX thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
XX or disorder, Type I diabetes, cancer, cell damage or damage to an
XX innocent bystander tissue that is induced by a chemotherapeutic agent or
XX therapeutic irradiation, treating haematopoietic progenitor cells that
XX have been exposed to therapeutic radiation or chemotherapy, aplastic
XX anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
XX also used for promoting the growth or differentiation of a haematopoietic
XX progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
XX resulting from abnormal myocardial ischaemia. The present sequence
XX represents human mFLINT.

XX Sequence 271 AA;

Query Match 69.3%; Score 671; DB 20; Length 271;
Best Local Similarity 100.0%; Pred. No. 9e-50;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 VAETPTVWRDAETGERLVCAQCPPTGVQRPCRRDSPTCGPPRHVYTFQWNYLRCR 89
DB 1 vaetptywrdaetgerlvcaqcpptgvqpcrrdspmtcgprrhytqfwnylrcr 60
QY 90 YCNVLCGEREEERACHATHNRCRCRTGFFAHAGFCLHASCPPGAGVIAPG 142
DB 61 ycnvlgereeeearachathnrcrctgffahagfclhascppgagviapg 113

RESULT 14

Y42185
ID Y42185 standard; Protein: 273 AA.

XX Y42185;

XX 17-DEC-1999 (first entry)

XX Human mFLINT #2 protein sequence.

XX Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;

Db 1 lercrycnvlgereeearachathnracrcrtgffahagfclehascpvgagviapg 58

Search completed: January 30, 2001, 16:45:27
Job time: 587 sec

A:Title: Molecular cloning and expression of the type 1 and type 2 murine receptors for
A:Reference number: A40254; MUID:91246168
A:Accession: A40254
A:Molecule type: mRNA
A:Residues: 1-474 <GOO>
A:CROSS-references: GB:M60469; NID:g199827; PIDN:AAA39752.1; PID:g199828
R:Kissonerghis, M.; Fellows, R.; Feldmann, M.; Chernajovsky, Y.
submitted to the EMBL Data Library, May 1995
A:Description: Characterization of the promoter region of the murine p75-TNF receptor.
A:Reference number: S54816
A:Accession: S54816
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-22 <KTS>
A:CROSS-references: EMBL:X87128; NID:g809043; PIDN:CAA60618.1; PID:g809044
C:Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology
F:Keywords: cytokine receptor; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-474/Product: tumor necrosis factor receptor type 2 #status predicted <MAT>
F:40-77/Domain: NGF receptor repeat homology <NG1>
F:79-120/Domain: NGF receptor repeat homology <NG2>
F:166-203/Domain: NGF receptor repeat homology <NG4>

Query Match 21.0%; Score 203; DB 2; Length 474;
Best Local Similarity 31.8%; Pred. No. 1.7e-09;
Matches 42; Conservative 13; Mismatches 55; Indels 22; Gaps 3;

QY 46 RLVAQCPPTFTVORPCRRDPTTCGPPRHVTQFNWYLERCYNVLCGEREEARAC 105
: : ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : : |||||
Db 52 QMCAKCPGQYKHFCKNTSDTVACDEASMTQVMNQFTCLSSSSCTTQDQVEIRAC 111
: : ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : : |||||

QY 106 HATHNACRRTGFF----AHAGF--CLHASCPPGAGVIAPGESWARGAPRSGCRR 158
: : ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : : |||||
Db 112 TKQNRVCAACEAGRYCALKTHSGSCRCMLSKGPGFGV-----ASSRA 156
: : ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : : |||||

QY 159 GRQVAGPSLAP 170
: : ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : : |||||
Db 157 PNGNVLCACAP 168
: : ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : : |||||

RESULT 3
A35356
tumor necrosis factor receptor type 2 precursor - human
N:Alternate names: 75K tumor necrosis factor receptor
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A35356; A36475; A48416; A36007; A32666; B35010; I38094
R:Smith, C.A.; Davis, T.; Anderson, D.; Solam, L.; Beckmann, M.P.; Jerzy, R.; Dower, S.K.
Science 248, 1019-1023, 1990
A:Title: A receptor for tumor necrosis factor defines an unusual family of cellular and
A:Reference number: A35356; MUID:90260639
A:Accession: A35356
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-461 <SMI>
A:CROSS-references: GB:M32315; NID:g189185; PIDN:AAA59929.1; PID:g189186
R:Kohn, T.; Brewer, S.L.; Schwartz, P.E.; King, M.W.; Hale, K.K.; Squires,
Proc. Natl. Acad. Sci. U.S.A. 87, 8331-8335, 1990
A:Title: A second tumor necrosis factor receptor gene product can shed a naturally occur
A:Reference number: A36475; MUID:91045991
A:Accession: A36475
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-195 'R', 197-461 <KOH>
A:CROSS-references: GB:M55994; GB:M38549; NID:g339757; PIDN:AAA36755.1; PID:g339758
R:Dembic, Z.; Loetscher, H.; Gubler, U.; Pan, Y.C.; Lahm, H.W.; Gentz, R.; Brockhaus, M.
Cytokine 2, 231-237, 1990
A:Title: Two human TNF receptors have similar extracellular, but distinct intracellular,
A:Reference number: A48416; MUID:91370690
A:Accession: A48416
A>Status: preliminary
A:Molecule type: mRNA; protein

A:Residues: 23-461 <DEM>
A:CROSS-references: GB:S63368; NID:g235648; PIDN:AAAB19824.1; PID:g235649
A>Note: sequence extracted from NCBI backbone (NCBIN:63368, NCBIIP:63371)
R:Heller, R.A.; Song, K.; Onasch, M.A.; Fischer, W.H.; Chang, D.; Ringold, G.M.
Proc. Natl. Acad. Sci. U.S.A. 87, 6151-6155, 1990
A:Title: Complementary DNA cloning of a receptor for tumor necrosis factor and demons
A:Reference number: A36007; MUID:90349572
A:Accession: A36007
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 116-140, 'P', 142-195, 'R', 197-362, 'T', 364-461 <HEL>
A:CROSS-references: GB:M35857; NID:g339751; PIDN:AAA63562.1; PID:g339752
R:Loetscher, H.; Schlaeger, E.J.; Lahm, H.W.; Pan, Y.C.E.; Lesslauer, W.; Brockhaus,
J. Biol. Chem. 265, 20131-20138, 1990
A:Title: Purification and partial amino acid sequence analysis of two distinct tumor
A:Reference number: A32666; MUID:91056048
A:Accession: A32666
A>Status: preliminary
A:Molecule type: protein
A:Residues: 23-40; 65-69; 136-141; 300-306 <LOE>
R:Engelmann, H.; Novick, D.; Wallach, D.
J. Biol. Chem. 265, 15311-15316, 1990
A:Title: Two tumor necrosis factor-binding proteins purified from human urine. Eviden
A:Reference number: A35010; MUID:90110215
A:Accession: B35010
A>Status: preliminary
A:Molecule type: protein
A:Residues: 27-31 <ENG>
R:Kuhnert, P.; Kemper, O.; Wallach, D.
Gene 150, 381-386, 1994
A:Title: Cloning, sequencing and partial functional characterization of the 5' region
A:Reference number: I38094; MUID:95121934
A:Accession: I38094
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-37 <RES>
A:CROSS-references: EMBL:X80021; NID:g666044; PIDN:CAA56324.1; PID:g825701
C:Genetics:
A:Gene: GDB:TNFR2
A:CROSS-references: GDB:I25914; OMIM:191191
A:Map position: lp36.2-lp36.2
A:Introns: 26/3
A>Note: the list of introns is incomplete
C:Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology
C:Keywords: duplication; glycoprotein; receptor; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-416/Product: tumor necrosis factor receptor type 2 #status experimental <MAT>
F:40-76/Domain: NGF receptor repeat homology <NG1>
F:78-119/Domain: NGF receptor repeat homology <NG2>
F:120-162/Domain: NGF receptor repeat homology <NG3>
F:164-201/Domain: NGF receptor repeat homology <NG4>
F:262-279/Domain: transmembrane #status predicted <TMN>
F:280-461/Domain: intracellular #status predicted <INT>
F:171.193/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 20.9%; Score 202; DB 1; Length 461;
Best Local Similarity 31.4%; Pred. No. 2.1e-09;
Matches 48; Conservative 12; Mismatches 63; Indels 30; Gaps 4;

QY 8 GLSELCLVLPALLPVPVARGVAETPTVWRDAETGE-----RLVCAOCPPG 55
: : ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : : |||||
Db 13 GLELWAAHALPA-----QVAFTPYAP-----EPGSTRRLREYDQTAQMCCKSPG 60
: : ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : : |||||

QY 56 TFVORPCRRDPTTCGPPRHVTQFNWYLERCYNVLCGEREEARACHATHNACRC 115
: : ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : : |||||
Db 61 QHAKVCTKTSIDVDCSDSTYTQLWNWPECLSCGSCSSDQVETQACTREQNRICTC 120
: : ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : : |||||

QY 116 RTGFFAHAG-----FCLHASCPPGAGVIAPG 142
: : ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : : |||||
Db 121 RPGWYCALSKQEGCRCLCAPLRKCRPGFGVARPG 153
: : ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : : |||||

A:Title: Potential virulence determinants in terminal regions of variola smallpox virus
A:Reference number: Z20488; MUID:94088747
A:Accession: T28623
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-348 <DNA>
A:Cross-references: EMBL:L22579; NID:G623595; PIDN:AAA60933.1; PID:G439102
A:Experimental source: strain Bangladesh 1975
C:Superfamily: myxoma virus T2 protein; NGF receptor repeat homology

Query Match 20.0%; Score 193.5; DB 2; Length 348;
Best Local Similarity 31.4%; Pred. No. 8.1e-09;
Matches 43; Conservative 14; Mismatches 65; Indels 15; Gaps 3;
QY 9 LSLCLVLPALLPVPVAVRGVAETPTYPWRDAE-TGERLVCAQCPPGTFVORPCRRDSP 67
DB 10 LFLUSCIINGRAAPYTP-----PNGCKDTEYKRHLCCCLSCPPGTYSRLCDSKTN 61
QY 68 TTTCGCPPPRHVYTFQWNYLRCRYCNVLCGEREEARACHATHNRACRRTGFF-----A 121
DB 62 TQCTPGSGTFTSRNHLPAFLSCNCRNSQVETPSCTNTHNRICECSPGYYCLLKGS 121
QY 122 HAGFCLHASCPPGAGV 138
DB 122 GCKACVSQTKCGIGYGV 138

RESULT 8
B43692
T2 protein - rabbit fibroma virus
C:Species: rabbit fibroma virus, Shope fibroma virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: B43692
R:Opton, C.; DeLange, A.M.; McFadden, G.
Virology 160, 20-30, 1987
A:Title: Tumorigenic poxviruses: genomic organization and DNA sequence of the telomeric
A:Reference number: A43692; MUID:87321103
A:Accession: B43692
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-325 <DPT>
A:Cross-references: GB:M17433
C:Superfamily: myxoma virus T2 protein; NGF receptor repeat homology
F:64-105/Domain: NGF receptor repeat homology <NG2>
F:106-147/Domain: NGF receptor repeat homology <NG3>

Query Match 17.8%; Score 172.5; DB 2; Length 325;
Best Local Similarity 27.9%; Pred. No. 4e-07;
Matches 43; Conservative 14; Mismatches 76; Indels 21; Gaps 3;
QY 11 LLLCLVLPALLPVPVAVRGVAETPTYPWRDAETGERLVCAQCPPGTFVORPCRRDSP 70
DB 8 LVCVYVYGDVPSYQKCGGHDY-----EKDGLCCASCHPGFYASRLCGFSNTVC 61
QY 71 GPCCPPRHVYTFQWNYLRCRYCNVLCGEREEARACHATHNRACRRTGFFA-----HAG 124
DB 62 SPEDGTFTASTNHAPACVSCRCPTGHLSESPQCDRTHRCVNCSTGNYCLLKQNGCR 121
QY 125 FCLHASCPPGAGVIAPGESWARGGAPRSGGRR 158
DB 122 ICAPQTKCPAGYGV-----SGHTRAGDTLC 146

RESULT 9
GQVZML
T2 protein - myxoma virus (strain Lausanne)
C:Species: myxoma virus
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 18-Jun-1999
C:Accession: A40566
R:Opton, C.; Macen, J.L.; Schreiber, M.; McFadden, G.
Virology 184, 370-382, 1991

A:Title: Myxoma virus expresses a secreted protein with homology to the tumor necrosis
A:Reference number: A40566; MUID:91335768
A:Accession: A40566
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-326 <UPT>
A:Cross-references: GB:M95181; GB:M37976; NID:G332309; PIDN:AAA46632.1; PID:G9323310
C:Superfamily: myxoma virus T2 protein; NGF receptor repeat homology
C:Keywords: glycoprotein
F:64-105/Domain: NGF receptor repeat homology <NG2>
F:106-147/Domain: NGF receptor repeat homology <NG3>
F:66.181.205.238/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 17.6%; Score 170; DB 1; Length 326;
Best Local Similarity 28.5%; Pred. No. 6.3e-07;
Matches 45; Conservative 13; Mismatches 74; Indels 26; Gaps 4;
QY 12 LCLVLPALL-----PVPVAVRGVAETPTYPWRDAETGERLVCAQCPPGTFVORPCRRDS 66
DB 4 LTLAYVACVYGGAPYGADRKCRGNDY-----EKDGLCCTSCPPGSYASRLCGPGS 57
QY 67 PTTCCGCPPPRHVYTFQWNYLRCRYCNVLCGEREEARACHATHNRACRRTGFFA----- 121
DB 58 DTVCSPCKNETFTASTNHAPACVSCRCPTGHLSESPQCDKTRDRVCDSCAGNYCLLKQ 117
QY 122 -HAGFCLHASCPPGAGVIAPGESWARGGAPRSGGRR 158
DB 118 EGCRCAPKTKCPAGYGV-----SGHTRGDTLC 146

RESULT 10
I37552
OX40 homolog - human
C:Species: Homo sapiens (man)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 11-Jan-2000
C:Accession: I37552
R:Latz, U.; Durkop, H.; Schnittger, S.; Ringeling, J.; Eitelbach, F.; Hummel, M.; Fo
Eur. J. Immunol. 24, 677-683, 1994
A:Title: The human OX40 homolog: cDNA structure, expression and chromosomal assignment
A:Reference number: I37552; MUID:94170844
A:Accession: I37552
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-277 <RES>
A:Cross-references: EMBL:X75962; NID:G472957; PIDN:CAA53576.1; PID:G472958
C:Superfamily: CD27 antigen; NGF receptor repeat homology

Query Match 16.3%; Score 157.5; DB 2; Length 277;
Best Local Similarity 31.0%; Pred. No. 5.8e-06;
Matches 45; Conservative 17; Mismatches 64; Indels 19; Gaps 6;
QY 6 GPCSLCLLVLPALLPVPVAVRGVAETPTYPWRDAETGERLVCAQCPPGTFVORPCRRD 65
DB 11 GPACALLGLLGLSTVTGLHCY-----GDTYPSNDR-----CCHCECPNGMGVSCRS 59
QY 66 SPTTCGCPPPRHVYTFQWNY--LERCRCNVLCGEREEARACHATHNRACRRTGFFA 123
DB 60 QNTVRCPCGPGYNDVVSKPKCTWCNLRSG--SERKQLCTATQDTCRCRAGTQPLD 117
QY 124 GF--CLEHASCPPGAGVIAPGESWA 146
DB 118 SYKPGVDCAPCPPGH--FSPGDNQA 140

RESULT 11
A46517
CD27 antigen precursor - human
N:Alternate names: CD27L receptor; T cell activation antigen CD27
C:Species: Homo sapiens (man)
C:Date: 18-Jun-1993 #sequence_revision 22-Apr-1995 #text_change 22-Jun-1999
C:Accession: A46517; A46454
R:Loenen, W.A.; Gravesteln, L.A.; Beumer, S.; Mellef, C.J.; Hagemeljer, A.; Borst, J.

A:Accession: B40254
A:Molecule type: mRNA
A:Residues: 1-454 <GO2>
A:Cross-references: GB:M60468; NID:g199825; PIDN:AAA39751.1; PID:g199826
R:Barrett, K.; Taylor-Fishwick, D.A.; Cope, A.P.; Kissenerghis, A.M.; Gray, P.W.; Fel
Eur. J. Immunol. 21, 1649-1656, 1991
A:Title: Cloning, expression and cross-linking analysis of the murine p55 tumor necro
A:Reference number: S16677; MUID:91285014
A:Accession: S16677
A:Molecule type: mRNA
A:Residues: 1-454 <BAR>
A:Cross-references: EMBL:X59238; NID:g53578; PIDN:CAA41922.1; PID:g53579
R:Rothe, J.G.; Brockhaus, M.; Gentz, R.; Lesslauer, W.
Immunogenetics 34, 338-340, 1991
A:Title: Molecular cloning and expression of the mouse Tnf receptor type b.
A:Reference number: S19021; MUID:92039815
A:Accession: S19021
A:Molecule type: mRNA
A:Residues: 1-454 <ROT>
A:Cross-references: EMBL:X57796; NID:g54848; PIDN:CAA40936.1; PID:g54849
R:Bebo, B.F.
Immunogenetics 39, 450-451, 1994
A:Title: Nucleotide sequence of the TNF type I receptor from a mouse endothelioma cel
A:Reference number: I54532; MUID:94245292
A:Accession: I54532
A>Status: translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-454 <RES>
A:Cross-references: GB:L26349; NID:g430732; PIDN:AAAS9361.1; PID:g430733
C:Comment: This protein is one of two distantly related receptors for both TNF-alpha
C:Superfamily: tumor necrosis factor receptor type 1; NGF receptor repeat homology
C:Keywords: duplication; glycoprotein; receptor; transmembrane protein
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-454/Product: tumor necrosis factor receptor type 1 #status predicted <MAT>
F:30-212/Domain: extracellular #status predicted <EXT>
F:84-82/Domain: NGF receptor repeat homology <NG1>
F:84-126/Domain: NGF receptor repeat homology <NG2>
F:127-167/Domain: NGF receptor repeat homology <NG3>
F:168-204/Domain: NGF receptor repeat homology <NG4>
F:213-235/Domain: transmembrane #status predicted <MEM>
F:236-454/Domain: intracellular #status predicted <INT>
F:54,151,202/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 13.7%; Score 132.5; DB 1; Length 454;
Best Local Similarity 22.2%; Pred. No. 0.00093;
Matches 46; Conservative 20; Mismatches 68; Indels 73; Gaps 7;

Qy 7 PGLSLLCLVLALPALLPVAVRGVAETPTVWRDA-----ETGERLVCAQCQP 54
||| : ||| : | : | : | :
Db 7 PGLLSLVLLALLGIIHPSGVTGL--VPSLGDKREKDSLCPQGKYVHSHNSICTTKCHK 64
||| : ||| : | : | : | :
Qy 55 GTTFVORPCRDRSP---TTCGPCPPRHVTQFNWLYLCRYC-----N 92
||| : ||| : | : | : | :
Db 65 GTYLVSDC--PSPGRDVCRECEKGSTFTASQNYLRQLCSCKTKCKMSQVEI5PCQAQDK 122
||| : ||| : | : | : | :
Qy 93 VLCCEREERAR-----ACHATHNRACRCRTGFFAHAGFCLEH 129
||| : ||| : | : | : | :
Db 123 TVCCCKENQRFQYLSETHFQCVDSCPFNGTVITPKCTQTVCNCIHAGFFLRESECVPC 182
||| : ||| : | : | : | :
Qy 130 ASC-----PPCAGVIAPGES 144
||| : ||| : | : | : | :
Db 183 SHCKKNECMKLCUPPLANVTNPQDS 209
||| : ||| : | : | : | :

RESULT 13
I57826
tumor necrosis factor receptor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 23-Jul-1999
C:Accession: I57826
R:Rothe, J.G.; Bluethmann, H.; Gentz, R.; Lesslauer, W.; Steinmetz, M.
Mol. Immunol. 30, 165-176, 1993

Search completed: January 30, 2001, 16:50:32
Job time: 722 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 17:06:06 ; Search time 80.04 seconds
(without alignments)
68.591 Million cell updates/sec

Title: US-09-518-931-4
Perfect score: 968
Sequence: 1 MRALGCPGLSLCLVLALPA.....PRSGRRRCRGVAGPSLAP 170

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	203	21.0	474	1 TNR2_MOUSE	P25119 mus musculus
2	202	20.9	461	1 TNR2_HUMAN	P20333 homo sapien
3	201.5	20.8	435	1 TNR2_HUMAN	P36941 homo sapien
4	194	20.0	349	1 VC22_VARV	P34015 variola vir
5	172.5	17.8	325	1 VT2_SEVKA	P25943 shope fibro
6	170	17.6	326	1 VT2_MYXVL	P29825 myxoma viru
7	168.5	17.4	415	1 TNR2_MOUSE	P50284 mus musculus
8	157.5	16.3	277	1 OX40_HUMAN	P43489 homo sapien
9	139.5	14.4	260	1 CD27_HUMAN	P26842 homo sapien
10	137.5	14.2	471	1 TNR1_BOVIN	O19131 bos taurus
11	133	13.7	332	1 FASA_PIG	O77736 sus scrofa
12	132.5	13.7	454	1 TNR1_MOUSE	P25118 mus musculus
13	132	13.6	272	1 OX40_MOUSE	P47741 mus musculus
14	131.5	13.6	271	1 OX40_RAT	P15725 rattus norv
15	128	13.2	595	1 CD30_HUMAN	P28908 homo sapien
16	127	13.1	323	1 FASA_BOVIN	P51867 bos taurus
17	127	13.1	324	1 FASA_MOUSE	P25446 mus musculus
18	125.5	13.0	324	1 FASA_RAT	O63199 rattus norv
19	122	12.6	256	1 41BB_MOUSE	P20334 mus musculus
20	122	12.6	335	1 FASA_HUMAN	P25445 homo sapien
21	120	12.4	277	1 CD40_HUMAN	P25942 homo sapien
22	118.5	12.2	255	1 41BB_HUMAN	OQ7011 homo sapien
23	117.5	12.1	250	1 CD27_MOUSE	P41272 mus musculus
24	117	12.1	269	1 CD40_BOVIN	Q28203 bos taurus
25	116	12.0	461	1 TNR1_RAT	P22934 rattus norv
26	111.5	11.5	416	1 NGFR_CHICK	P18519 gallus gall
27	110	11.4	289	1 CD40_MOUSE	P27512 mus musculus
28	108.5	11.2	455	1 TNR1_HUMAN	P19438 homo sapien
29	106.5	11.0	425	1 NGFR_RAT	P07174 rattus norv
30	106	11.0	427	1 NGFR_HUMAN	P08138 homo sapien
31	102.5	10.6	461	1 TNR1_PIG	P50555 sus scrofa
32	98	10.1	727	1 META_DROME	Q24567 drosophila
33	94	9.7	915	1 PAC6_RAT	P41413 rattus norv

ALIGNMENTS

RESULT 1

ID	TNR2_MOUSE	STANDARD;	PRT;	474 AA.
AC	P25119; P97893;			
DT	01-MAY-1992 (Rel. 22, Created)			
DT	01-MAY-1992 (Rel. 22, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	TUMOR NECROSIS FACTOR RECEPTOR 2 PRECURSOR (TNF-R2) (P75).			
GN	TNFRSF1B OR TNFR2 OR TNFR-2.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91187885; PubMed=1849278;			
RA	Lewis M., Tartaglia L.A., Lee A., Bennett G.L., Rice G.C.,			
RA	Wong G.H., Chen E.Y., Goeddel D.V.;			
RT	"Cloning and expression of cDNAs for two distinct murine tumor			
RT	necrosis factor receptors demonstrate one receptor is species			
RT	specific."			
RL	Proc. Natl. Acad. Sci. U.S.A. 88:2830-2834(1991).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91246168; PubMed=1645445;			
RA	Goodwin R.G., Anderson D., Jerzy T., Davis T., Brannan C.I.,			
RA	Copeland N.G., Jenkins N.A., Smith C.A.;			
RT	"Molecular cloning and expression of the type 1 and type 2 murine			
RT	receptors for tumor necrosis factor."			
RL	Mol. Cell. Biol. 11:3020-3026(1991).			
RN	[3]			
RP	SEQUENCE OF 1-26 FROM N.A.			
RC	STRAIN=NOD;			
RA	Jacob C.O., Liu J.;			
RA	Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE OF 1-22 FROM N.A.			
RC	TISSUE=LIVER;			
RA	Kisssonerthis M., Fellows R., Feldmann M., Chernajovsky Y.;			
RA	Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.			
CC	-!- FUNCTION: RECEPTOR FOR TNF-ALPHA.			
CC	-!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.			
CC	-!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; M60469; AAA39752.1; -			
DR	EMBL; M59378; AAA40463.1; -			
DR	EMBL; X59488; AAA85021.1; -			
DR	EMBL; X87128; CAA60618.1; -			
DR	PIR; B38634; B38634.			

P31695 mus musculus
O05793 mus musculus
Q28983 sus scrofa
P55268 homo sapien
Q93038 h wsl-1 pro
P04592 mus musculus
P18128 escherichia
P35442 homo sapien
O05895 mus musculus
P35448 xenopus lae
P21071 vaccinia vi
P24756 vaccinia vi

34 94 9.7 1964 1 NTC4_MOUSE
35 91.5 9.5 3707 1 PGBM_MOUSE
36 91 9.4 2476 1 ZAN_PIG
37 90.5 9.3 1798 1 LMB2_HUMAN
38 90 9.3 417 1 WSL1_HUMAN
39 90 9.3 915 1 PAC6_MOUSE
40 88 9.1 245 1 YPC2_ECOLI
41 88 9.1 1172 1 TSP2_HUMAN
42 87 9.0 956 1 TSP3_MOUSE
43 87 9.0 1173 1 TSP1_XENLA
44 86 8.9 103 1 VA53_VACCC
45 86 8.9 103 1 VA53_VACCV

Kohno T., Brewer M.T., Baker S.L., Schwartz P.E., King M.W.,
Hale K.K., Squires C.H., Thompson R.C., Vannice J.L.;
"A second tumor necrosis factor receptor gene product can shed a
naturally occurring tumor necrosis factor inhibitor.";
Proc. Natl. Acad. Sci. U.S.A. 87:8331-8335(1990).
[3]
RN SEQUENCE FROM N.A.
RX MEDLINE=96299745; PubMed=8661109;
RA Beltinger C.P., White P.S., Maris J.M., Sulman E.P., Jensen S.J.,
Lepassier D., Stallard B.J., Goeddel D.V., Desauvage F.J.,
Rodeur G.M.;
RT "Physical mapping and genomic structure of the human TNFR2 gene.";
RL Genomics 35:94-100(1996).
[4]
RN SEQUENCE OF 116-461 FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=90349572; PubMed=2165946;
RA Heller R.A., Song K., Onasch M.A., Fischer W.H., Chang D.,
Ringold G.M.;
RT "Complementary DNA cloning of a receptor for tumor necrosis factor
and demonstration of a shed form of the receptor.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:6151-6155(1990).
[5]
RN SEQUENCE OF 27-31.
RX MEDLINE=90110215; PubMed=2153136;
RA Engelmann H., Novick D., Wallach D.;
RT "Two tumor necrosis factor-binding proteins purified from human
urine. Evidence for immunological cross-reactivity with cell surface
tumor necrosis factor receptors";
RL J. Biol. Chem. 265:1531-1536(1990).
[6]
RN SEQUENCE OF 22-40; 65-69; 136-141; 300-306 AND 346-362.
RX MEDLINE=91056048; PubMed=2173696;
RA Loetscher H., Schlaeger E.J., Lahm H.-W., Pan Y.-C.E., Lesslauer W.,
Brockhaus M.;
RT "Purification and partial amino acid sequence analysis of two
distinct tumor necrosis factor receptors from HL60 cells.";
RL J. Biol. Chem. 265:20131-20138(1990).
[7]
RN CHARACTERIZATION.
RX MEDLINE=93016040; PubMed=1328224;
RA Pennica D., Lam V.T., Mize N.K., Weber R.F., Fendly B.M.,
Lipari M.T., Goeddel D.V.;
RT "Biochemical properties of the 75-kDa tumor necrosis factor receptor.
Characterization of ligand binding, internalization, and receptor
phosphorylation.";
RL J. Biol. Chem. 267:21172-21178(1992).
[8]
RN X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 419-428 IN TRAF2 COMPLEX.
RX MEDLINE=99221490; PubMed=10206649;
RA Park Y.C., Burkitt V., Villa A.R., Tong L., Wu H.;
RT "Structural basis for self-association and receptor recognition of
human TRAF2.";
RL Nature 398:533-538(1999).
CC -!- FUNCTION: RECEPTOR FOR TNF-ALPHA. HIGH AFFINITY FOR TNA-ALPHA AND
APPROXIMATELY 5-FOLD LOWER AFFINITY FOR TNF-BETA.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- PTM: PHOSPHORYLATED; MAINLY ON SERINE RESIDUES WITH A VERY LOW
LEVEL ON THREONINE RESIDUES.
CC -!- PHARMACEUTICAL: AVAILABLE UNDER THE NAME ENBREL (IMMUNEX AND
WYETH-AVERTIS). USED TO TREAT MODERATE TO SEVERE RHEUMATOID
ARTHRITIS (RA). ENBREL CONSIST OF THE EXTRACELLULAR LIGAND-BINDING
PORTION OF TNFR2 LINKED TO AN IMMUGLOBULIN FC CHAIN. IT BINDS TO
TNF-ALPHA AND BLOCKS ITS INTERACTIONS WITH RECEPTORS.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD120b entry;
WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd120b.htm"
CC -!- DATABASE: NAME=Enbreil; NOTE=Clinical information on Enbreil;
WWW="http://www.enbrelinfo.com/"

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>) or send an email to license@isb-sib.ch).

CC EMBL; M32315; AAA59929.1; -;
DR EMBL; M35857; AAB63262.1; -;
DR EMBL; U52165; AAC50622.1; JOINED.
DR EMBL; U52156; AAC50622.1; JOINED.
DR EMBL; U52157; AAC50622.1; JOINED.
DR EMBL; U52158; AAC50622.1; JOINED.
DR EMBL; U52159; AAC50622.1; JOINED.
DR EMBL; U52160; AAC50622.1; JOINED.
DR EMBL; U52161; AAC50622.1; JOINED.
DR EMBL; U52162; AAC50622.1; JOINED.
DR EMBL; U52163; AAC50622.1; JOINED.
DR EMBL; U52164; AAC50622.1; JOINED.
DR EMBL; M55994; AAA36755.1; -;
DR PIR; A35356; A35356.
DR PIR; A36007; A36007.
DR PIR; A36475; A36475.
DR PIR; B35010; B35010.
DR PIR; A23666; A23666.
DR PDB; 1CA9; 12-APR-99.
DR MIM; 191191; -;
DR INTERPRO; IPR001368; -;
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 4.
DR KKK Receptor; Transmembrane; Glycoprotein; Repeat; Signal;
Phosphorylation; Pharmaceutical; 3D-structure.
CHAIN 1 22
DOMAIN 23 461 TUMOR NECROSIS FACTOR RECEPTOR 2.
TRANSMEM 23 257 EXTRACELLULAR (POTENTIAL).
CYTOPLASMIC 258 287 POTENTIAL.
DOMAIN 288 461 CYTOPLASMIC (POTENTIAL).
DOMAIN 39 201 4 X TNFR-CYS.
REPEAT 39 76 TNFR-CYS 1.
REPEAT 77 118 TNFR-CYS 2.
REPEAT 119 162 TNFR-CYS 3.
REPEAT 163 201 TNFR-CYS 4.
DISULFID 40 53 BY SIMILARITY.
DISULFID 54 67 BY SIMILARITY.
DISULFID 57 75 BY SIMILARITY.
DISULFID 78 93 BY SIMILARITY.
DISULFID 96 110 BY SIMILARITY.
DISULFID 100 118 BY SIMILARITY.
DISULFID 120 126 BY SIMILARITY.
DISULFID 134 143 BY SIMILARITY.
DISULFID 137 161 BY SIMILARITY.
DISULFID 164 179 BY SIMILARITY.
CARBOHYD 171 171 N-LINKED (GLCNAC. . .) (POTENTIAL).
CARBOHYD 193 193 N-LINKED (GLCNAC. . .) (POTENTIAL).
CONFLICT 141 141 R -> P (IN REF. 4).
CONFLICT 196 196 R -> M (IN REF. 1 AND 3).
CONFLICT 363 363 A -> T (IN REF. 4).
SEQUENCE 461 AA; 48316 MW; 603B580ECD67636F CRC64;

Query Match 20.98; Score 202; DB 1; Length 461;
Best Local Similarity 31.4%; Pred. No. 3.6e-10;
Matches 48; Conservative 12; Mismatches 63; Indels 30; Gaps 4;

QY 8 GLSLCLVLALPVARGVAETPTYPWRDAETGE-----RLVCAQCPPG 55
| | | | | | | | | | | | | | | : : : | | |
Db 13 GLELWAAHALPA-----QVAFTPVP----EPGSTCRLEYDYDTQMCKSCSPG 60
| | | | | | | | | | | | | | | : : : | | |
QY 56 TEVQPCRRDSPPTTCGCPPHRYTFWNLYBRCRYCNVLCGEREEEAACHATHNRACR 115
| | | | | | | | | | | | | | | : : : | | |
Db 61 QHAKVFCTKTDVTCDSCEDSTYTQLNWNVEPELSCSGRCSSDQVETQACTREQNRICTC 120
| | | | | | | | | | | | | | | : : : | | |
QY 116 RFGFFAHAG-----FCLEASHACPPGGAGVIAPG 142
| | | | | | | | | | | | | | | : : : | | |
Db 121 RFGWCALSKOEGRCRLCAPLRKRPGFGVARPG 153
| | | | | | | | | | | | | | | : : : | | |

RESULT 3
TNRC_HUMAN STANDARD; PRT; 435 AA.
ID TNRC_HUMAN AC P36941.
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE LYMPHOTOXIN-BETA RECEPTOR PRECURSOR (TUMOR NECROSIS FACTOR RECEPTOR)
DE 2 RELATED PROTEIN) (TUMOR NECROSIS FACTOR C RECEPTOR).
GN LTBR OR TNFR OR TNFRSF3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER.
RX MEDLINE=93252381; PubMed=8486360;
RA Baens M., Chaffanet M., Cassiman J.J., den Berghe H., Marynen P.;
"Construction and evaluation of a hncDNA library of human 12p
transcribed sequences derived from a somatic cell hybrid.";
RL Genomics 16:214-218(1993).
RN [2]
RP FUNCTION.
RX MEDLINE=94225209; PubMed=8171323;
RA Crowe P.D., van Arsdale T.L., Walter B.N., Ware C.F., Hession C.,
RA Ehrenfels B., Browning J.L., Din W.S., Goodwin R.G., Smith C.A.;
"A lymphotoxin-beta-specific receptor.";
RL Science 264:707-710(1994).
CC -!- FUNCTION: RECEPTOR FOR THE LYMPHOTOXIN-BETA. POSSIBLE FUNCTION IN
IMMUNE DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: CONTAINS A LA-NGR/TNFR-TYPE CYSTEINE-RICH REGION.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed, usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
or send an email to license@isb-sib.ch.

CC EMBL; L04270; AAA36757.1; -;
DR HSSP; P25942; 1CDF.
DR MIM; 600979; -;
DR INTERPRO; IPR001368; -;
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 3.
KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 30 POTENTIAL.
FT CHAIN 31 435 LYMPHOTOXIN-BETA RECEPTOR.
FT DOMAIN 31 227 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 228 248 POTENTIAL.
FT DOMAIN 249 435 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 42 211 4 X TNFR-CYS.
FT REPEAT 42 81 TNFR-CYS 1.
FT REPEAT 82 124 TNFR-CYS 2.
FT REPEAT 125 168 TNFR-CYS 3.
FT REPEAT 169 211 TNFR-CYS 4.
FT DISULFID 43 58 BY SIMILARITY.
FT DISULFID 59 72 BY SIMILARITY.
FT DISULFID 62 80 BY SIMILARITY.
FT DISULFID 83 98 BY SIMILARITY.
FT DISULFID 101 116 BY SIMILARITY.
FT DISULFID 104 124 BY SIMILARITY.
FT DISULFID 126 132 BY SIMILARITY.
FT DISULFID 139 148 BY SIMILARITY.
FT DISULFID 142 167 BY SIMILARITY.
FT DISULFID 170 185 BY SIMILARITY.
FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 435 AA; 46709 MW; 624626E6022F656F CRC64;
Query Match 20.8%; Score 201.5; DB 1; Length 435;
Best Local Similarity 35.5%; Pred. No. 3.8e-10;
Matches 54; Conservative 14; Mismatches 59; Indels 25; Gaps 8;
QY 3 ALGGPGLSLCLVLPALLPVPVAVGVAETPTV-----PWRDA-----ETGERLVCAQC 52
DB 6 ATSAPGLWGLPLVGLFGLLAASQPAV---PPYASENOTCRDQEKYEPQHRICCSRC 62
QY 53 PPGTFVQRPQRDRSPPTGCPGPPRHHTQWNYL---EGRYCNVLCGEREEARACHATH 109
DB 63 PPGTYVSASKSRIRDTVCATCAENSYNEHWNYLTICQLCRPCDPVNG--LEETAPCTSKR 120
QY 110 NRACRCRTGFAHAGFCLE--H---ASCPPG 135
DB 121 KTCRCQPGMFC-AWALECTHCELLSDCPG 151
RESULT 4
VC22_VARV STANDARD; PRT; 349 AA.
AC P34015;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE PROTEIN C22/B28 HOMOLOG.
GN G4R.
OS Variola virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-INDIA-1967 / ISOLATE IND3;
RX MEDLINE-93202281; PubMed-8384129;
RA Shchelkunov S.N., Blinov V.M., Sandakchchiev L.S.;
RT "Genes of variola and vaccinia viruses necessary to overcome the host protective mechanisms.";
RL FEBS Lett. 319:80-83(1993).
CC -!- SIMILARITY: CONTAINS 2 LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGIONS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC
CC EMBL; X69198; CAA49137.1; -
CC DR EMBL; X67117; CAA47540.1; -
CC DR PIR; D36858; D36858.
CC DR PIR; S35987; S35987.
CC DR PIR; S46888; S46888.
CC DR HSSP; P19438; INCF.
CC DR INTERPRO: IPR001368; -.
CC DR PFAM: PF00020; TNFR_C6; 2.
CC DR PROSITE; PS00652; TNFR_NGFR_1; 2.
CC DR PROSITE; PS50050; TNFR_NGFR_2; 2.
CC DR PROSITE; PS50050; TNFR_NGFR_2; 2.
CC KW REPEAT.
CC FT DOMAIN 31 108 2 X TNFR-CYS.
CC FT REPEAT 31 66 TNFR-CYS 1.
CC FT REPEAT 67 108 TNFR-CYS 2.
CC FT REPEAT 349 AA; 38189 MW; D45D40B5C6E780EF CRC64;
SQ SEQUENCE 349 AA; 38189 MW; D45D40B5C6E780EF CRC64;
Query Match 20.0%; Score 194; DB 1; Length 349;
Best Local Similarity 30.1%; Pred. No. 1.3e-09;
Matches 41; Conservative 15; Mismatches 68; Indels 12; Gaps 2;
QY 9 LSLCLVLVALPALLPVPVAVGVAETPTVPRDAGTGERLVCAQCPTGTVQRPQRDRSPT 68

DB 10 LFLSCIIINGROAAPVTPPNKCKDTEY-----KRHNLCCLSCPPGYGYASRLDCKNTNT 63
QY 69 TCGPCPPRHHTQFWNYLRCRYCNVLCGEREEARACHATHNRACRGTGFE-----AH 122
DB 64 QCTPCGSGTFTSRNNHLPACLSCNGRCNSNQVETRSCNTHNRICECSPGYYCLLKSSG 123
QY 123 AGFCLEHASCPPGAGV 138
DB 124 CKACVSOCTKCGIGYV 139
RESULT 5
VT2_SFVKA STANDARD; PRT; 325 AA.
AC P25943;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE TUMOR NECROSIS FACTOR SOLUBLE RECEPTOR PRECURSOR (PROTEIN T2).
GN T2.
OS Shope fibroma virus (strain Kasza) (SFV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Leporipoxvirus.
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE-87321103; PubMed-2820128;
RA Upton C., Delange A.M., McFadden G.;
RT "Tumorigenic poxviruses: genomic organization and DNA sequence of the telomeric region of the Shope fibroma virus genome.";
RL Virology 160:20-30(1987).
RN [2]
RP FUNCTION.
RX MEDLINE-91207415; PubMed-1850261;
RA Smith C.A., Davis T., Wignall J.M., Din W.S., Farrah T., Upton C., McFadden G., Goodwin R.G.;
RT "T2 open reading frame from the Shope fibroma virus encodes a soluble form of the TNF receptor.";
RL Biochem. Biophys. Res. Commun. 176:335-342(1991).
CC -!- FUNCTION: BINDS TO TNF-ALPHA AND BETA. PROBABLY PREVENTS TNF TO REACH CELLULAR TARGET AND THEREBY DEAMPENING THE POTENTIAL ANTIVIRAL EFFECTS OF THE CYTOKINE.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC
CC EMBL; M17433; -; NOT_ANNOTATED_CDS.
CC DR EMBL; A23727; CAA01687.1; -
CC DR PIR; B43692; B43692.
CC DR HSSP; P19438; ITNR.
CC DR INTERPRO: IPR001368; -.
CC DR PFAM: PF00020; TNFR_C6; 2.
CC DR PROSITE; PS00652; TNFR_NGFR_1; 2.
CC DR PROSITE; PS00652; TNFR_NGFR_2; 1.
CC DR PROSITE; PS50050; TNFR_NGFR_2; 1.
CC KW Receptor; Glycoprotein; Repeat; Signal.
CC FT CHAIN 1 16 POTENTIAL.
CC FT DOMAIN 17 325 TUMOR NECROSIS FACTOR SOLUBLE RECEPTOR.
CC FT REPEAT 27 186 4 X TNFR-CYS.
CC FT REPEAT 27 62 TNFR-CYS 1.
CC FT REPEAT 63 104 TNFR-CYS 2.
CC FT REPEAT 105 147 TNFR-CYS 3.
CC FT REPEAT 148 186 TNFR-CYS 4.
CC FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT SEQUENCE 325 AA; 35132 MW; 81053039198A71E CRC64;
SQ SEQUENCE 325 AA; 35132 MW; 81053039198A71E CRC64;

Query Match	17.6%	Score 170;	DB 1;	Length 326;
Best Local Similarity	28.5%;	Pred. No. 1.3e-07;		
Matches	45;	Conservative 13;	Mismatches 74;	Indels 26; Gaps 4;
QY	12	LCILVLALPALL----	PVPAVRGVAETPTYPWRDAETGERLVCACQPPGTFVQPCRRDS	66
Db				
QY	4	LTLLLAYVACVYGGGAPYGADRKGCRNDY-----	EKGDLCCITSCPPTSASRLCQPGS	57
Db				
QY	67	PTTCGCPGPRHYTFQWNYLERCYCNVLCGEREBEARACHATHNRACRCRTGFFA-----		121
Db				
QY	58	DTVCSPECKNETFTASTNHAPACVSCRCRCCTHLSSESQCDKTRDVRDCDSAGNVCLLKQG		117
Db				
QY	122	-HAGFCLEHASCPPGAGVIAPGESWARGGAPRSGRRRC		158
Db				
QY	118	EGCRICAPKTKCPAGYGV-----	SGHRTGTGDLVC	146
Db				

```

RESULT 7
ID      TNRC_MOUSE      STANDARD;      PRT;      415 AA.
TN      AC
DT      P50284;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      01-NOV-1997 (Rel. 35, Last annotation update)
DE      LYMPHOTOXIN-BETA RECEPTOR PRECURSOR.
GN      LTBR OR TNFCR.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN      [1]
RC      SEQUENCE FROM N.A.
RC      STRAIN-CVB; TISSUE-LUNG;
RX      MEDLINE-96072804; PubMed-7594541;
RA      Force W.R., Walter B.N., Hession C., Tizard R., Kozak C.A.,
RA      Browning J.L., Ware C.F.;
RT      "Mouse lymphotoxin-beta receptor. Molecular genetics, ligand binding,
RT      and expression." ;
RL      J. Immunol. 155:5280-5288(1995).
RN      [2]
RC      SEQUENCE FROM N.A.
RC      MEDLINE-96163885; PubMed-8586432;
RA      Nakamura T., Tashiro K., Nazarea M., Nakano T., Sasayama S.,
RA      Honjo T.;
RA      "The murine lymphotoxin-beta receptor cDNA: isolation by the signal
RT      sequence trap and chromosomal mapping." ;
RL      Genomics 30:312-319(1995).
CC      -1- FUNCTION: RECEPTOR FOR THE LYMPHOTOXIN-BETA. POSSIBLE FUNCTION IN
CC      IMMUNE DEVELOPMENT.
CC      -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN
CC      -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC      -----
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; U29173; AAA68964.1; -
CC      DR      EMBL; L38423; AAB00846.1; -
CC      DR      EMBL; U30798; AAB81334.1; -
CC      DR      HSSP; P25942; ICDF.
CC      DR      MGD; MGI:104875; LTBR.
CC      DR      INTERPRO; IPR001368; -
CC      DR      PFAM; PF00020; TNFR_C6; 3.
CC      DR      PROSITE; PS00652; TNFR_NGFR_1; 2.
CC      DR      PROSITE; PS50050; TNFR_NGFR_2; 3.
CC      DR      Receptor; Transmembrane; Glycoprotein; Repeat; Signal.
KW      SIGNAL
FT      CHAIN          1      30      POTENTIAL.
FT      DOMAIN          31      415      LYMPHOTOXIN-BETA RECEPTOR.
FT      DOMAIN          31      223      EXTRACELLULAR (POTENTIAL).

```

```

CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X75962; CAA53576.1;
CC EMBL; S76792; AAB33944.1; ALT_INIT.
CC HSP; P25942; ICDF.
CC MIN; 600315;
CC INTERPRO; IPR001368;
CC PFAM; PF00020; TNFR_C6; 3.
CC PROSITE; PS00652; TNFR_NGFR_1; 3.
CC PROSITE; PS00652; TNFR_NGFR_2; 2.
CC Receptor; T-cell; Antigen; Glycoprotein; Transmembrane; Repeat;
KW SIGNAL.
FT CHAIN 1 28
FT DOMAIN 29 277
FT DOMAIN 29 214
FT TRANSMEM 215 235
FT DOMAIN 236 277
FT DOMAIN 30 169
FT REPEAT 30 65
FT REPEAT 66 107
FT REPEAT 108 126
FT REPEAT 127 167
FT CARBOHYD 146 146
FT CARBOHYD 160 160
FT SEQUENCE 277 AA; 29340 MW; 49F15259411550BF CRC64;

Query Match 16.3%; Score 157.5; DB 1; Length 277;
Best Local Similarity 31.0%; Pred. No. 1.2e-06;
Matches 45; Conservative 17; Mismatches 64; Indels 19; Gaps 6;

QY 6 GPGSLLLCLVLPVAVRGVAETPTYPWRDAETGERLVCAQCPPTGVQPCRRD 65
DB 11 GPCAAALLGLGLSTVTGLHCY-----GDTYPSNDR-----CCHCECRPGNGMWSRCSRS 59

QY 66 SPTTCGCPPRHYTQFWNY--LERCYCNVLCGEREEARACHATHNRACRRTGFFAFA 123
DB 60 QNTVCRPCGPGFYNDVSVSKPKPCPTWCNLSRG--SERKQLCTATQDTCRCRAGTQPLD 117

QY 124 GF--CLEHASCPPGAGVIAPGESWA 146
DB 118 SYKPGVDCAPCPGPH--FSPGDNQA 140

RESULT 9
CD27_HUMAN
ID CD27_HUMAN STANDARD; PRT; 260 AA.
AC P26842;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CD27L RECEPTOR PRECURSOR (T-CELL ACTIVATION ANTIGEN CD27) (T14).
GN TNFRSF7 OR CD27.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MONOCYTES;
RX MEDLINE=92013149; PubMed=1655907;
RA Camerini D., Walz G., Loenen W.A.M., Borst J., Seed B.;
RT "The T cell activation antigen CD27 is a member of the nerve growth
RT factor/tumor necrosis factor receptor gene family.";
RL J. Immunol. 147:3165-3169(1991).
RN [2]
RP SEQUENCE FROM N.A.

```



```

DR DISULFID 182 190 BY SIMILARITY.
DR DISULFID 185 194 BY SIMILARITY.
FT CARBOHYD 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 145 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 471 AA; 51367 MW; 5243EF514DFE81C4 CRC64;

Query Match 14.2%; Score 137.5; DB 1; Length 471;
Best Local Similarity 31.6%; Pred. No. 9.2e-05;
Matches 49; Conservative 13; Mismatches 58; Indels 35; Gaps 10;

QY 7 PGLSLCLVLALPALLP-VPA-VRGVAETPTYPWRDAETGE-----RLVC 49
DB 7 PG-----LLPLVLALLADVYPAGVQGLVPHP-----GDLEKRESPPQKYNHPQNSTICC 59

QY 50 AQCPPTGTVORPC-----RRDPTTCGCPPPHYTFQFNLYLRCRYCNVLCGER-EPEARAC 105
DB 60 TKHKGTYLYNDPCPGPRD--TDCRCAPGTYYTALENHLRLCLSCRCRDEMFOVEISPC 117

QY 106 HATHNRACRCRTG-----FFAHAGF-CLEHASCPPG 135
DB 118 VDRDVTGCRKNQYREYWGTEGFRCLNCSLCPNG 152

RESULT 11
FASA_PIG STANDARD; PRT; 332 AA.
AC 07736;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE FASL RECEPTOR PRECURSOR (APOPTOSIS-MEDIATING SURFACE ANTIGEN FAS)
DE (APO-1 ANTIGEN) (CD95).
GN TNFRSF6 OR APTL OR FAS.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN [1]
RP Barling B., Hoffmann J., Holtz J., Schulz R., Heusch G., Darner D.;
RT "Expression of apoptosis-associated genes in hibernating and stunned
myocardium of pig."
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS FASL. THE
CC ADAPTOR MOLECULE FADD RECRUITS CASPASE-8 TO THE ACTIVATED
CC RECEPTOR. THE RESULTING AGGREGATE CALLED THE DEATH-INDUCING
CC SIGNALING COMPLEX (DISC) PERFORMS CASPASE-8 PROTEOLYTIC
CC ACTIVATION. ACTIVE CASPASE-8 INITIATES THE SUBSEQUENT CASCADE OF
CC CASPASES (ASPARTATE-SPECIFIC CYSTEINE PROTEASES) MEDIATING
CC APOPTOSIS. FAS-MEDIATED APOPTOSIS MAY HAVE A ROLE IN THE
CC INDUCTION OF PERIPHERAL TOLERANCE. IN THE ANTIGEN-STIMULATED
CC SUICIDE OF MATURE T-CELLS, OR BOTH (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN
CC -1- DOMAIN: CONTAINS A DEATH DOMAIN INVOLVED IN THE BINDING OF FADD,
CC AND MAYBE TO OTHER CYTOSOLIC ADAPTOR PROTEINS.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ001202; CAA04596.1; -
CC INTERPRO; IPR000488; -
CC DR INTERPRO; IPR001368; -
CC DR PFAM; PF00020; TNFR_c6; 3.
CC DR PFAM; PF00531; death; 1.
CC DR PROSITE; PS00652; TNFR_NGFR_1; 2.

```

```

DR PROSITE; PS50050; TNFR_NGFR_2; 2.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
KW Apoptosis; Receptor; Glycoprotein; Transmembrane; Repeat; Signal.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 332 FASL RECEPTOR.
FT DOMAIN 17 175 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 176 192 POTENTIAL.
FT DOMAIN 193 332 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 45 164 3 X TNFR-CYS.
FT REPEAT 45 81 TNFR-CYS 1.
FT REPEAT 82 125 TNFR-CYS 2.
FT REPEAT 126 164 TNFR-CYS 3.
FT DOMAIN 227 311 DEATH DOMAIN.
FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 332 AA; 37592 MW; 5B8B03682756BFB1B CRC64;

Query Match 13.7%; Score 133; DB 1; Length 332;
Best Local Similarity 34.6%; Pred. No. 0.00016;
Matches 37; Conservative 8; Mismatches 46; Indels 16; Gaps 7;

QY 39 RDAETGE-----RLVCAQCPPTGTFVQRCRRDSP-----TTCGCPPT-RHYTFQFNLYLRC 88
DB 42 RESECPGQHGREGFCPCQPCPPGKRKHADC--TSPGAPQCVPCSEGEDYTDKNHSSKC 99

QY 89 RYCNVLCGERE-EEARACHATHNRACRCRTGFFAHAGFCLHASCPP 134
DB 100 RRCRVCDGEGHGLEVEKNTQNTKCKCKPNFFCHTSQC-EH--CNP 143

RESULT 12
TNRL_MOUSE STANDARD; PRT; 454 AA.
AC P25118;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR 1 PRECURSOR (P60) (TNF-R1) (TNF-R1)
DE (P55).
GN TNFRSF1A OR TNFR1 OR TNFR-1.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RP MEDLINE-91187885; PubMed-1849278; Bennett G.L., Rice G.C.,
RA Lewis M., Tartaglia L.A., Lee A., Chen E.Y., Goeddel D.V.;
RT "Cloning and expression of cDNAs for two distinct murine tumor
RT necrosis factor receptors demonstrate one receptor is species
RT specific."
RT Proc. Natl. Acad. Sci. U.S.A. 88:2830-2834(1991).
RN [2]
RP SEQUENCE FROM N.A.
RP MEDLINE-91246168; PubMed-1645445;
RA Goodwin R.G., Anderson D., Jerzy R., Davis T., Brannan C.I.,
RA Copeland N.G., Jenkins N.A., Smith C.A.;
RT "Molecular cloning and expression of the type 1 and type 2 murine
RT receptors for tumor necrosis factor."
RT Mol. Cell. Biol. 11:3020-3026(1991).
RN [3]
RP SEQUENCE FROM N.A.
RP MEDLINE-91285014; PubMed-1647956;
RA Barrett K., Taylor-Fishwick D.A., Cope A.P., Kissonerghis A.M.,
RA Gray P.W., Feldmann M., Foxwell B.M.J.;
RT "Cloning, expression and cross-linking analysis of the murine p55
RT tumor necrosis factor receptor."
RT Eur. J. Immunol. 21:1649-1656(1991).
RN [4]
RP SEQUENCE FROM N.A.
RP TISSUE-SPLEEN.
RC MEDLINE-92039815; PubMed-1657766;
RX

```

Rothe J.G., Brockhaus M., Gentz R., Lesslauer W.;
 "Molecular cloning and expression of the mouse Tnf receptor type b.";
 Immunogenetics 34:338-340(1991).
 [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94245292; PubMed=8188324;
 Bebo B.F., Linthicum D.S.;
 "Nucleotide sequence of the TNF type I receptor from a mouse
 endothelioma cell line.";
 Immunogenetics 39:450-451(1994).
 [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93156721; PubMed=8381516;
 Rothe J., Bluetmann H., Gentz R., Lesslauer W., Steinmetz M.;
 "Genomic organization and promoter function of the murine tumor
 necrosis factor receptor beta gene.";
 Mol. Immunol. 30:165-175(1993).
 CC -!- FUNCTION: RECEPTOR FOR TNF-ALPHA. THE ADAPTOR MOLECULE FADD
 RECRUITS CASPASE-8 TO THE ACTIVATED RECEPTOR. THE RESULTING
 AGGREGATE CALLED THE DEATH-INDUCING SIGNALING COMPLEX (DISC)
 PERFORMS CASPASE-8 PROTEOLYTIC ACTIVATION WHICH INITIATES THE
 SUBSEQUENT CASCADE OF CASPASES (ASPARTATE-SPECIFIC CYSTEINE
 PROTEASES) MEDIATING APOPTOSIS (BY SIMILARITY).
 CC -!- SUBUNIT: TNF BINDING TO THE EXTRACELLULAR DOMAIN OF TNFR1 LEADS TO
 HOMOTRIMERIZATION. ONCE AGGREGATED THE RECEPTORS DEATH DOMAINS
 PROVIDE A NOVEL MOLECULAR INTERFACE THAT INTERACTS SPECIFICALLY
 WITH THE DEATH DOMAIN OF TRADD. VARIOUS TRADD-INTERACTING
 PROTEINS SUCH AS TRAFs, RIP AND POSSIBLY FADD, ARE RECRUITED TO
 TNFR1 COMPLEX BY THEIR ASSOCIATION WITH TRADD. THIS COMPLEX
 ACTIVATES AT LEAST TWO DISTINCT SIGNALING CASCADES, APOPTOSIS AND
 NF-KAPPA B SIGNALING (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -!- SIMILARITY: CONTAINS A LA-NGR/TNFR-TYPE CYSTEINE-RICH REGION.
 CC -!- SIMILARITY: CONTAINS 1 DEATH DOMAIN.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; M60468; AAA39751.1; -
 DR EMBL; M59377; AAA0464.1; -
 DR EMBL; X52338; CAA41922.1; -
 DR EMBL; X57796; CAA40936.1; -
 DR EMBL; L26349; AAA59361.1; -
 DR EMBL; M76656; AAA40465.1; -
 DR EMBL; M88067; AAA40465.1; JOINED.
 DR EMBL; M76655; AAA40465.1; JOINED.
 DR PIR; A38634; GQMS1.
 DR PIR; S16677; S16677.
 DR PIR; S19021; S19021.
 DR HSP; P19438; LEXT.
 DR MGD; MGI:1314884; TNFRSF1A.
 DR INTERPRO: IPR000488; -
 DR INTERPRO: IPR001368; -
 DR PRAM; PF00020; TNFR_c6; 4.
 DR PRAM; PF00531; death; 1.
 DR PROSITE; PS00652; TNFR_NGFR_1; 3.
 DR PROSITE; PS50050; TNFR_NGFR_2; 3.
 DR PROSITE; PS50017; DEATH_DOMAIN; 1.
 KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal; Apoptosis.
 FT SIGNAL 1 21
 FT CHAIN 22 454
 FT DOMAIN 22 212 TUMOR NECROSIS FACTOR RECEPTOR 1.
 FT TRANSMEM 213 235 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 236 454 POTENTIAL.
 FT DOMAIN 43 196 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 43 196 4 X TNFR-CYS.
 FT REPEAT 83 125 TNFR-CYS 1.
 FT REPEAT 83 125 TNFR-CYS 2.
 FT REPEAT 126 166 TNFR-CYS 3.

FT REPEAT 167 196 TNFR-CYS 4.
 FT DOMAIN 339 349 N-SMASE ACTIVATION DOMAIN (NSD).
 FT DOMAIN 356 441 DEATH DOMAIN.
 FT DISULFID 44 58 BY SIMILARITY.
 FT DISULFID 59 72 BY SIMILARITY.
 FT DISULFID 62 81 BY SIMILARITY.
 FT DISULFID 84 99 BY SIMILARITY.
 FT DISULFID 102 117 BY SIMILARITY.
 FT DISULFID 105 125 BY SIMILARITY.
 FT DISULFID 127 143 BY SIMILARITY.
 FT DISULFID 146 158 BY SIMILARITY.
 FT DISULFID 149 166 BY SIMILARITY.
 FT DISULFID 168 179 BY SIMILARITY.
 FT DISULFID 182 191 BY SIMILARITY.
 FT DISULFID 185 195 BY SIMILARITY.
 FT CARBOHYD 54 54 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 151 151 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 202 202 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CONFLICT 394 394 R -> G (IN REF. 6).
 SQ SEQUENCE 454 AA; 50129 MW; 0710C2E8C3C2B6D9 CRC64;
 Query Match 13.7%; Score 132.5; DB 1; Length 454;
 Best Local Similarity 22.2%; Pred. No. 0.00023;
 Matches 46; Conservative 20; Mismatches 68; Indels 73; Gaps 7;
 QY 7 PGLSLCLVLALPALLPVPVAVGVAEPTYPWRDA-----ETGERLYCAQCAPP 54
 DB 7 PGLLSLVLLALLGHIHPGVTGL--VPSLGDREKROSLCPGKYVYKSNNSICTRCHK 64
 QY 55 GTFVQRPCRDRSP---TTCGCPRRHYTFQWNYLERCRYC-----N 92
 DB 65 GTYLVSDC--PSPGRDVTVCRECEKGTFTASONYLRQCLSKCTCKEMSOVLSPCQADKD 122
 QY 93 VLGEREEEAR-----ACHATHNRACRCRTGTFFAHAGFCLLEH 129
 DB 123 TVCGCKENQFQRYLSETHFQCVDCSPCFNGVTPTPKETQNTVCNCHAGFFLRESECVPC 182
 QY 130 ASC-----PGCAGVIAFGES 144
 DB 183 SHCKKNECMKMLCPPPLANYTNPQDS 209
 RESULT 13
 OX40_MOUSE
 ID OX40_MOUSE STANDARD; PRT; 272 AA.
 AC P47741;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE OX40L RECEPTOR PRECURSOR (OX40 ANTIGEN).
 GN TNFRSF4 OR TXGP1 OR OX40.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/C;
 RX MEDLINE=94044750; PubMed=8228223;
 RA Calderhead D.M., Buhlmann J.E., van den Eertwegh A.J.,
 RA Claassen E., Noelle R.J., Fell H.;
 RT "Cloning of mouse OX40: a T cell activation marker that may mediate
 T-B cell interactions";
 RL J. Immunol. 151:5261-5271(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95255413; PubMed=7737295;
 RA Birkeland M.L., Copeland N.G., Gilbert D.J., Jenkins N.A.,
 RA Barclay A.N.;
 RT "Gene structure and chromosomal localization of the mouse homologue
 of rat OX40 protein.";
 RL Eur. J. Immunol. 25:926-930(1995).
 CC -!- FUNCTION: RECEPTOR FOR THE OX40L/GP34 CYTOKINE.

```

C      -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
C      -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RIK REGION.
C      -----
C      This SWISS-PROT entry is copyright. It is produced through a collaboration
C      between the Swiss Institute of Bioinformatics and the EMBL outstation -
C      the European Bioinformatics Institute. There are no restrictions on its
C      use by non-profit institutions as long as its content is in no way
C      modified and this statement is not removed. Usage by and for commercial
C      entities requires a license agreement (See http://www.isb-sib.ch/announcement/
C      or send an email to license@isb-sib.ch).
C      -----
C      EMBL: Z21674; CAA79772.1; -
C      EMBL: X85214; CAA59476.1; -
C      HSP: P25942; ICDF.
C      MGI: MGI:104512; TXGPI.
C      INTERPRO: IPR001368; -
C      PFAM: PF00020; TNFR_C6; 3.
C      PROSITE: PS00652; TNFR_NGFR_1; 3.
C      PROSITE: PS50050; TNFR_NGFR_2; 2.
C      Receptor; T-cell; Antigen; Glycoprotein; Transmembrane; Repeat;
C      Signal.
C      SIGNAL 1 19 POTENTIAL.
C      CHAIN 20 272 OX40L RECEPTOR.
C      DOMAIN 20 211 EXTRACELLULAR (POTENTIAL).
C      TRANSEM 212 236 POTENTIAL.
C      DOMAIN 237 272 CYTOPLASMIC (POTENTIAL).
C      DOMAIN 26 165 4 X TNFR-CYS.
C      REPEAT 26 61 TNFR-CYS 1.
C      REPEAT 62 103 TNFR-CYS 2. (INCOMPLETE).
C      REPEAT 104 124 TNFR-CYS 3.
C      REPEAT 125 165 TNFR-CYS 4.
C      CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
C      CONFLICT 15 15 A -> G (IN REF. 2).
C      SEQUENCE 272 AA; 30153 MW; 06E7BB4156F0D08E CRC64;
C
QY      Query Match 13.6%; Score 132; DB 1; Length 272;
      Best Local Similarity 31.2%; Pred. No. 0.00017;
      Matches 43; Conservative 15; Mismatches 64; Indels 16; Gaps 7
C
QY      15 VLALPALLVPAVGVAEITPTVPRDAETGERLVCAQCPGPTFVQRPGRDSPTTCGCP 74
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      11 LLLLATLGLVTARLLNCKVHTP-----SGHK-CCRECQPGHGMVSRCDHTDRLCHPCE 64
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      75 PRHYTFQWNY--LERCRYCNVLGGREREAREACHATHNRACRCRTGFAH--AGFCI--E 128
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      65 TGFYNEAVNYDTCKQTCQNRHRS--SELKNCCTPTQDTCVRCRPTQPRODSYKLGVD 122
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      129 HASCPPGAGVIAPGESWA 146
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      123 CVPCPPGH--FSPGNQQA 138
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
C
RESULT 14
OX40L_RAT STANDARD; PRT; 271 AA.
AC P15725;
C      01-APR-1990 (Rel. 14, Created)
C      01-APR-1990 (Rel. 14, Last sequence update)
C      01-OCT-2000 (Rel. 40, Last annotation update)
C      01-OCT-2000 (Rel. 40, Last annotation update)
C      OX40L RECEPTOR PRECURSOR (OX40 ANTIGEN) (MRC OX40).
C      TNFRSF4 OR TXGPIL OR OX40.
C      Rattus norvegicus (Rat).
C      OX40L Receptor; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
C      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
C      [1]
C      SEQUENCE FROM N.A.
C      RP TISSUE-T-CELL;
C      RX MEDLINE=90214614; PubMed=2157591;
C      RA Mallett S.; Fossum S.; Barclay A.N.;
C      RT "Characterization of the MRC OX40 antigen of activated CD4 positive T
C      RL lymphocytes -- a molecule related to nerve growth factor receptor.";
C      FMBO J. 9:1063-1068(1990).

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:54:55 ; Search time 254.27 Seconds
(without alignments)
78.363 Million cell updates/sec

Title: US-09-518-931-4
Perfect score: 968
Sequence: 1 MRALRGPLSLICLVLPALP.....PRSGRRGRCRGVAGPSLAP 170

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_15:*

- 1: sp.archaea:*
- 2: sp.bacteria:*
- 3: sp.fungi:*
- 4: sp.human:*
- 5: sp.invertebrate:*
- 6: sp.mammal:*
- 7: sp.mhc:*
- 8: sp.organelle:*
- 9: sp.phage:*
- 10: sp.plant:*
- 11: sp.rodent:*
- 12: sp.virus:*
- 13: sp.vertebrate:*
- 14: sp.unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	814	84.1	300	4 095407	095407 homo sapien
2	296	30.6	401	4 000300	000300 homo sapien
3	292.5	30.2	372	4 09UHP4	Q9UHP4 homo sapien
4	285.5	29.5	401	11 008712	008712 mus musculus
5	282.5	29.2	401	11 008727	008727 rattus norv
6	230.5	23.8	302	13 Q9PUS0	Q9PUS0 salvelinus
7	206	21.3	459	11 Q62327	Q62327 mus musculus
8	204	21.1	348	12 057277	057277 monkeypox v
9	204	21.1	349	12 057100	057100 monkeypox v
10	202	20.9	348	12 057103	057103 monkeypox v
11	202	20.9	348	12 057108	057108 monkeypox v
12	202	20.9	349	12 057291	057291 monkeypox v
13	202	20.9	349	12 057101	057101 monkeypox v
14	202	20.9	349	12 057102	057102 monkeypox v
15	201	20.8	349	12 057099	057099 monkeypox v
16	199.5	20.6	482	11 088734	088734 mus musculus
17	194	20.0	349	12 057284	057284 camelpox vi
18	194	20.0	349	12 057098	057098 camelpox vi
19	194	20.0	349	12 057110	057110 variola vir

20	194	20.0	349	12	057111	057111 variola vir
21	194	20.0	349	12	089118	089118 variola vir
22	194	20.0	349	12	089098	089098 variola vir
23	193.5	20.0	348	12	057112	057112 variola vir
24	193.5	20.0	348	12	085407	085407 variola vir
25	191	19.7	439	4	Q16042	Q16042 homo sapien
26	190	19.6	349	12	057097	057097 camelpox vi
27	190	19.6	351	12	057117	057117 cowpox viru
28	190	19.6	355	12	085308	085308 cowpox viru
29	189	19.5	351	12	073559	073559 cowpox viru
30	188.5	19.5	350	12	057116	057116 cowpox viru
31	188	19.4	326	12	057120	057120 cowpox viru
32	187.5	19.4	655	4	075509	075509 homo sapien
33	186	19.2	349	12	057109	057109 variola vir
34	185	19.1	326	12	057122	057122 cowpox viru
35	185	19.1	349	12	057305	057305 cowpox viru
36	184	19.0	350	12	057123	057123 cowpox viru
37	183	18.9	347	12	057119	057119 cowpox viru
38	183	18.9	360	12	057118	057118 cowpox viru
39	180	18.6	351	12	057121	057121 cowpox viru
40	179	18.5	347	12	057115	057115 cowpox viru
41	167	17.3	322	12	072761	072761 cowpox viru
42	166	17.1	320	12	057079	057079 cowpox viru
43	164	16.9	316	12	057092	057092 ectromelia
44	163	16.8	320	12	057091	057091 ectromelia
45	163	16.8	320	12	057300	057300 ectromelia

ALIGNMENTS

RESULT 1

095407 ID 095407 PRELIMINARY; PRT; 300 AA.

AC 095407;

DT 01-MAY-1999 (TrEMBLrel. 10, Created)

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)

DE DECOY RECEPTOR 3 (M68) (M68C) (M68E).

GN DCR3 OR TR6.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99087326; PubMed=9872321;

RA Pitti R.M., Marsters S.A., Lawrence D.A., Roy M., Kischkel F.C., Dowd P., Huang A., Donahue C.J., Sherwood S.W., Baldwin D.T., Godowski P.J., Wood W.I., Gurney A.L., Hillan K.J., Cohen R.L., Goddard A.D., Botstein D., Ashkenazi A.

RA "Genomic amplification of a decoy receptor for Fas ligand in lung and colon cancer."

RT Nature 396:699-703(1998).

RL [2]

RN SEQUENCE FROM N.A.

RC TISSUP-BLOOD;

RX MEDLINE=99253915; PubMed=10318773;

RA Yu K.Y., Kwon B., Ni J., Zhai Y., Ebner R., Kwon B.S.;

RT "A newly identified member of tumor necrosis factor receptor superfamily (TR6) suppresses LIGHT-mediated apoptosis."

RL J. Biol. Chem. 274:13733-13736(1999).

RN [3]

RP SEQUENCE FROM N.A.

RX TISSUP-PANCREAS;

RX MEDLINE=20122600; PubMed=10655513;

RA Bai C., Connolly B., Metzker M.L., Hilliard C.A., Liu X., Sandig V., Soderman A., Galloway S.M., Liu Q., Austin C.P., Caskey C.T.;

RT "Overexpression of M68/DCR3 in human gastrointestinal tract tumors independent of gene amplification and its location in a four-gene cluster."

RT Proc. Natl. Acad. Sci. U.S.A. 97:1230-1235(2000).

RL EMBL; AF104419; AAD03056.1; -.

[illegible]

Qy	34	PTYPWRDAETGERLVCAOC	PGFTFVQPRCRDSPTTCGCP	PRHYTQFWNYLRCRCNV	93
Db	26	PKYLVDPTEGHLQCLD	CAFGPTGLVKOBCTVRRKTL	CVPCDHSYDSDSHSTDECVCSP	85
Qy	94	LCGEREEERACHATHN	ACRCRTGFFAHAGFCLEHASC	PPGAGVIAPEGSMWAGAPRS	153
Db	86	VCKELQSVQECNRTNR	VRNVRCECEGRYLEIEFCL	KHRSCPPGSGV-----	QAGTPR 138
Qy	154	GG--RRCGRGVAG--	PSLAP 170		
		::			
Db	139	NTVCKKCPDGFSTG	SSKAP 159		
RESULT	5				
O08727					
ID	008727	PRELIMINARY:	PRT:	401 AA.	
AC	O08727:				
DT	01-JUL-1997	(TEMBLrel. 04, Created)			
DT	01-JUL-1997	(TEMBLrel. 04, Last sequence update)			
DT	01-OCT-2000	(TEMBLrel. 15, Last annotation update)			
DE	OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR) (OCIF).				
GN	TNFRSF11B OR OPG.				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_TaxID=10116;				
RN	[1]				
RP	SEQUENCE FROM N. A.				
RC	TISSUE=INTESTINE;				
RX	MEDLINE=97262071; PubMed=9108485;				
RA	Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,				
RA	Luethy R., Nguyen H.Q., Wooden S., Bennett L., Boone T., Shimamoto G.,				
RA	DeRose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,				
RA	Davey E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,				
RA	Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,				
RA	Suggs S., Boyle W.J.;				
RT	"Osteoprotegerin: a novel secreted protein involved in the regulation				
RT	of bone density.";				
RL	Cell 89:309-319(1997).				
CC	FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES				
CC	OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY				
CC	SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO				
CC	OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN				
CC	STROMAL CELLS AND OSTEOCLAST PROGENITORS				
CC	SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM) (BY				
CC	SIMILARITY).				
CC	SUBCELLULAR LOCATION: EXTRACELLULAR.				
CC	SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.				
DR	EMBL:	U94330;	AAB53707.1;	..	
DR	HSSP:	P25942;	1CDF		
DR	INTERPRO:	IPR001368;	..		
DR	PFAM:	PF00020;	TNFR_C6;	4.	
DR	PROSITE:	PS00652;	TNFR_NGFR_1;	UNKNOWN_1.	
DR	PROSITE:	PS00050;	TNFR_NGFR_2;	2.	
DR	PRODOM:	PD000771;	..	1.	
KW	Glycoprotein; Repeat; Cytokine; Signal.				
FT	SIGNAL	1	21	BY SIMILARITY.	
FT	CHAIN	22	401	OSTEOPROTEGERIN.	
FT	DOMAIN	23	201	4 X TNFR-CYS.	
FT	REPEAT	23	63	TNFR-CYS 1.	
FT	REPEAT	64	106	TNFR-CYS 2.	
FT	REPEAT	107	143	TNFR-CYS 3.	
FT	REPEAT	144	201	TNFR-CYS 4.	
FT	DOMAIN	306	365	DEATH DOMAIN.	
FT	DISULFID	41	54	BY SIMILARITY.	
FT	DISULFID	44	62	BY SIMILARITY.	
FT	DISULFID	65	80	BY SIMILARITY.	
FT	DISULFID	83	97	BY SIMILARITY.	</

69 TCGPCPPRHYYTQFWNYLRCRYCNVLCGEREEARACHATHNRACRGTGFF-----AH 122
 64 QCTPCGSDTFTSHNHLOACLSCNGRCDNSQVETRSCNTHNRICCSPGYYCLLKGS 123
 123 AGFCLEHASCPCPGAGV 138
 124 CRTCSKTKCGIGYV 139

RESULT 12
 057291 PRELIMINARY; PRT; 349 AA.
 ID 057291
 AC 057291
 DT 01-JUN-1998 (Tremblrel. 06, Created)
 DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
 GN CRMB.
 OS Monkeypox virus.
 OS Viruses: dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Orthopoxvirus.
 OC NCBI_TaxID=10244;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VARIOUS STRAINS:
 RA Loparev V.N., Parsons J.M., Esposito J.J.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U88144; AAB94369.1; -
 DR EMBL; U87842; AAB94359.1; -
 DR EMBL; U87994; AAB94365.1; -
 DR EMBL; U87995; AAB94366.1; -
 DR EMBL; U88143; AAB94368.1; -
 DR HSP; P25942; ICDF.
 DR INTERPRO: IPR001368; -
 DR PFAM: PF00020; TNFR_C6; 2.
 DR PROSITE: PS00652; TNFR_NGFR_1; 2.
 DR PROSITE: PS00050; TNFR_NGFR_2; 2.
 DR PRODOM: PD000771; -; 1.
 DR PRODOM: PD000771; -; 1.
 SQ SEQUENCE 349 AA; 38295 MW; CBD2C949ED2B8E7C CRC64;

Query Match 20.9%; Score 202; DB 12; Length 349;
 Best Local Similarity 30.9%; Pred. No. 4.8e-12;
 Matches 42; Conservative 18; Mismatches 64; Indels 12; Gaps 2;
 9 LSLCLVLALPALLPVPVAVGVAETPTYPWRDAETGERLVCAQCPCPGTFFVORPCRRDSPT 68
 10 LFLSCIIINGRDIAHPANGKCKNEYRSN-----LCCLSCPCPGTYASRLCDSKTNT 63
 69 TCGPCPPRHYYTQFWNYLRCRYCNVLCGEREEARACHATHNRACRGTGFF-----AH 122
 64 QCTPCGSDTFTSHNHLOACLSCNGRCDNSQVETRSCNTHNRICCSPGYYCLLKGS 123
 123 AGFCLEHASCPCPGAGV 138
 124 CRTCSKTKCGIGYV 139

RESULT 13
 057101 PRELIMINARY; PRT; 349 AA.
 ID 057101
 AC 057101
 DT 01-JUN-1998 (Tremblrel. 06, Created)
 DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
 GN CRMB.
 OS Monkeypox virus.
 OS Viruses: dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Orthopoxvirus.
 OC NCBI_TaxID=10244;
 RN [1]
 RP SEQUENCE FROM N.A.

GN Monkeypox virus.
 OS Viruses: dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Orthopoxvirus.
 OC NCBI_TaxID=10244;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ZAIRE-1979 (79-0005);
 RA Loparev V.N., Parsons J.M., Esposito J.J.;
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U87847; AAB94364.1; -
 DR HSP; P25942; ICDF.
 DR INTERPRO: IPR001368; -
 DR PFAM: PF00020; TNFR_C6; 2.
 DR PROSITE: PS00652; TNFR_NGFR_1; 2.
 DR PROSITE: PS00050; TNFR_NGFR_2; 2.
 DR PRODOM: PD000771; -; 1.
 DR PRODOM: PD000771; -; 1.
 SQ SEQUENCE 348 AA; 38184 MW; 34A5E66B27907B5 CRC64;

Query Match 20.9%; Score 202; DB 12; Length 348;
 Best Local Similarity 30.9%; Pred. No. 4.8e-12;
 Matches 42; Conservative 18; Mismatches 64; Indels 12; Gaps 2;
 9 LSLCLVLALPALLPVPVAVGVAETPTYPWRDAETGERLVCAQCPCPGTFFVORPCRRDSPT 68
 10 LFLSCIIINGRDIAHPANGKCKNEYRSN-----LCCLSCPCPGTYASRLCDSKTNT 63
 69 TCGPCPPRHYYTQFWNYLRCRYCNVLCGEREEARACHATHNRACRGTGFF-----AH 122
 64 QCTPCGSDTFTSHNHLOACLSCNGRCDNSQVETRSCNTHNRICCSPGYYCLLKGS 123
 123 AGFCLEHASCPCPGAGV 138
 124 CRTCSKTKCGIGYV 139

RESULT 11
 057108 PRELIMINARY; PRT; 348 AA.
 ID 057108
 AC 057108
 DT 01-JUN-1998 (Tremblrel. 06, Created)
 DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
 GN CRMB.
 OS Monkeypox virus.
 OS Viruses: dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Orthopoxvirus.
 OC NCBI_TaxID=10244;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ZAIRE-1970 (CONGO-8);
 RA Loparev V.N., Parsons J.M., Esposito J.J.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U88142; AAB94367.1; -
 DR HSP; P25942; ICDF.
 DR INTERPRO: IPR001368; -
 DR PFAM: PF00020; TNFR_C6; 2.
 DR PROSITE: PS00652; TNFR_NGFR_1; 2.
 DR PROSITE: PS00050; TNFR_NGFR_2; 2.
 DR PRODOM: PD000771; -; 1.
 DR PRODOM: PD000771; -; 1.
 SQ SEQUENCE 348 AA; 38212 MW; E555979057DEC91F CRC64;

Query Match 20.9%; Score 202; DB 12; Length 348;
 Best Local Similarity 30.9%; Pred. No. 4.8e-12;
 Matches 42; Conservative 18; Mismatches 64; Indels 12; Gaps 2;
 9 LSLCLVLALPALLPVPVAVGVAETPTYPWRDAETGERLVCAQCPCPGTFFVORPCRRDSPT 68
 10 LFLSCIIINGRDIAHPANGKCKNEYRSN-----LCCLSCPCPGTYASRLCDSKTNT 63

124 CBTCISYKTCCTCVCH 120

us-09-518-931-4.rspt

Wed Jan 31 13:44:47 2001

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:45:27 ; Search time 183.67 Seconds
(without alignments)
26.064 Million cell updates/sec

Title: US-09-518-931-4_COPY_31_170

Perfect score: 821

Sequence: 1 AETPTYPWRDAETGERLVCA.....PRSGGRCRGQVAGPSLAP 140

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq_36: *
- 1: /SIDSL1/gcgdata/geneseq/geneseq/AA1980.DAT:*
 - 2: /SIDSL1/gcgdata/geneseq/geneseq/AA1981.DAT:*
 - 3: /SIDSL1/gcgdata/geneseq/geneseq/AA1982.DAT:*
 - 4: /SIDSL1/gcgdata/geneseq/geneseq/AA1983.DAT:*
 - 5: /SIDSL1/gcgdata/geneseq/geneseq/AA1984.DAT:*
 - 6: /SIDSL1/gcgdata/geneseq/geneseq/AA1985.DAT:*
 - 7: /SIDSL1/gcgdata/geneseq/geneseq/AA1986.DAT:*
 - 8: /SIDSL1/gcgdata/geneseq/geneseq/AA1987.DAT:*
 - 9: /SIDSL1/gcgdata/geneseq/geneseq/AA1988.DAT:*
 - 10: /SIDSL1/gcgdata/geneseq/geneseq/AA1989.DAT:*
 - 11: /SIDSL1/gcgdata/geneseq/geneseq/AA1990.DAT:*
 - 12: /SIDSL1/gcgdata/geneseq/geneseq/AA1991.DAT:*
 - 13: /SIDSL1/gcgdata/geneseq/geneseq/AA1992.DAT:*
 - 14: /SIDSL1/gcgdata/geneseq/geneseq/AA1993.DAT:*
 - 15: /SIDSL1/gcgdata/geneseq/geneseq/AA1994.DAT:*
 - 16: /SIDSL1/gcgdata/geneseq/geneseq/AA1995.DAT:*
 - 17: /SIDSL1/gcgdata/geneseq/geneseq/AA1996.DAT:*
 - 18: /SIDSL1/gcgdata/geneseq/geneseq/AA1997.DAT:*
 - 19: /SIDSL1/gcgdata/geneseq/geneseq/AA1998.DAT:*
 - 20: /SIDSL1/gcgdata/geneseq/geneseq/AA1999.DAT:*
 - 21: /SIDSL1/gcgdata/geneseq/geneseq/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	821	100.0	170	19 W63623	Human tumour necro
2	667	81.2	245	20 Y28449	A human tumour nec
3	667	81.2	271	20 Y42184	Human mFLINT #1 pr
4	667	81.2	300	19 W66102	Amino acid sequenc
5	667	81.2	300	19 W63622	Human tumour necro
6	667	81.2	300	20 Y03099	Human lung TNF-rec
7	667	81.2	300	20 Y42182	Human FLINT #1 pro
8	667	81.2	300	20 Y17479	Mammalian tumour n
9	667	81.2	300	20 Y06817	Human DCR3 polypep
10	667	81.2	300	20 W97749	Human tumour necro
11	667	81.2	300	20 W95082	Orphan receptor (H
12	656	79.9	273	20 Y42185	Human mFLINT #2 pr

13	656	79.9	302	20 Y42183	Human FLINT #2 pro
14	653	79.5	300	21 Y77458	Human TNF receptor
15	337	41.0	153	20 Y22222	Human TNFR superfa
16	333	40.6	215	20 W93585	Human HAP06 protei
17	296.5	36.1	359	17 R99939	Mutated OCIF, OCIF
18	293.5	35.7	401	17 R99933	Mutated OCIF, OCIF
19	292.5	35.6	187	17 R99950	Mutated OCIF, OCIF
20	292.5	35.6	187	21 Y77464	Primate protein se
21	292.5	35.6	197	17 R99945	Mutated OCIF, OCIF
22	292.5	35.6	272	17 R99944	Mutated OCIF, OCIF
23	292.5	35.6	321	17 R99949	Mutated OCIF, OCIF
24	292.5	35.6	326	17 R99940	Mutated OCIF, OCIF
25	292.5	35.6	327	17 R99941	Mutated OCIF, OCIF
26	292.5	35.6	349	20 W83928	Human FTHMA-070 pa
27	292.5	35.6	351	17 R99943	Mutated OCIF, OCIF
28	292.5	35.6	380	17 R99924	Mature osteoclasto
29	292.5	35.6	390	17 R99357	Human tumour necro
30	292.5	35.6	391	19 W53238	Human OCIF genome
31	292.5	35.6	393	17 R99948	Mutated OCIF, OCIF
32	292.5	35.6	395	19 W57636	Modified TR1 recep
33	292.5	35.6	399	17 R99942	Mutated OCIF, OCIF
34	292.5	35.6	401	17 R99932	Mutated OCIF, OCIF
35	292.5	35.6	401	17 R99934	Mutated OCIF, OCIF
36	292.5	35.6	401	17 R99925	Full length osteoc
37	292.5	35.6	401	17 R99931	Mutated OCIF, OCIF
38	292.5	35.6	401	18 W38345	Human osteoprotege
39	292.5	35.6	401	19 W57635	TR1 receptor prote
40	292.5	35.6	401	19 W53239	Human OCIF genome
41	292.5	35.6	401	20 Y05742	Tumour necrosis fa
42	292.5	35.6	401	20 W95030	Tumour necrosis fa
43	292.5	35.6	401	20 W83926	Human FTHMA-070 pr
44	292.5	35.6	401	21 Y43400	Osteoprotegrin pro
45	292	35.6	143	17 R99946	Mutated OCIF, OCIF

ALIGNMENTS

RESULT 1

W63623 ID W63623 standard; Protein; 170 AA.

XX AC W63623;

XX DT 26-OCF-1998 (first entry)

XX DE Human tumour necrosis factor receptor-6 beta protein.

XX KW Human tumour necrosis factor receptor-6 alpha; TNFR-6 alpha; TNFR-6 beta; endothelial cells; keratinocytes; normal prostate; apoptosis; prostate tumour tissue.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Peptide 1..30

FT Protein 31..170

FT note="TNFR-6 beta"

FT Region 31..166

XX FT /note= "Soluble extracellular domain"

XX PN W09830694-A2.

XX PD 16-JUL-1998.

XX PF 13-JAN-1998; 98WO-US00153.

XX PR 14-JAN-1997; 97US-0035496.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Ebner R, Feng P, Gentz RL, Ni J, Ruben SM, Yu G;

DR	WPI; 1998-399142/34.
RR	N-PSDB; V39086.
XX	Human tumour necrosis factor receptors 6-alpha and 6-beta - used in
XX	the diagnosis of immune system-related disorder(s)
XX	Claim 20; Fig 2A-2B; 91pp; English.
XX	The present sequence represents the human tumour necrosis factor
CC	receptor-6 beta (TNFR-6 beta) protein. The invention also provides
CC	for the TNFR-6 alpha protein (W53622). TNFR-6 alpha and TNFR-6 beta are
CC	members of the tumour necrosis factor receptor (TNFR) family. TNFRs
CC	are expressed in endothelial cells, keratinocytes, normal prostate and
CC	prostate tumour tissue. For a number of disorders of these cells,
CC	particularly of the immune system, substantially altered (whether
CC	increased or decreased) levels of TNFR-6 alpha and/or TNFR-6 beta gene
CC	expression can be detected, therefore the TNFR-6 alpha and TNFR-6 beta
CC	polypeptides, nucleic acids and antibodies are claimed to be useful in
CC	the diagnosis of such disorders. Mutations of the TNFR-6 alpha and
CC	TNFR-6 beta genes can also be detected. The TNFR polypeptides are
CC	also claimed to be useful for identifying ligands which may be useful
CC	in the treatment of apoptosis related disorders.
XX	Sequence 170 AA;
XX	SQ
XX	Query Match 100.0%; Score 821; DB 19; Length 170;
XX	Best Local Similarity 100.0%; Pred. No. 3.1e-65;
XX	Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 AETPTYPWRDAETGERLVCAQCPCGTFVQRCDSPTCGCPPRHYTQFWNYLCRCY 60
DQ	
DB	31 AETPTYWRDAETGERLVCAQCPCGTFVQRCDSPTCGCPPRHYTQFWNYLCRCY 90
QY	61 CNVLGGEREEARACHATHNRACRRTGFFAHAGFCLEHASCPPGAGVIAPGSWARGGA 120
DB	
QY	121 PRSGGRRCRGVAGPSTLAP 140
DB	
QY	151 PRSGGRRCRGVAGPSTLAP 170
DB	
XX	RESULT 2
XX	ID Y28449 standard; Protein; 245 AA.
XX	AC Y28449;
XX	DT 29-SEP-1999 (first entry)
XX	A human tumour necrosis factor-R2-like proteins (TR2P)-1.
XX	Human tumour necrosis factor-R2-like protein; TR2P; achondroplasia;
KW	osteoporosis; developmental disorder; Cushing's syndrome;
KW	muscular dystrophy; epilepsy; hereditary neuropathy;
KW	Charcot-Marie-Tooth disease; neurofibromatosis; hypothyroidism;
KW	hydrocephalus; seizure disorder; cerebellar palsy; spinal bifida;
KW	congenital glaucoma; cataract; sensorineural hearing loss;
KW	reproductive disorder; infertility; ovulatory defect; endometriosis;
KW	autoimmune disorder; ectopic pregnancy; teratogenesis; spermatogenesis;
KW	immunological disorder; AIDS; Addison's disease; allergy; bronchitis;
KW	atherosclerosis; diabetes mellitus; Chron's disease; lupus;
KW	irritable bowel syndrome; multiple sclerosis; infection;
KW	neoplastic disorder; adenocarcinoma; leukaemia; lymphoma; melanoma;
KW	myeloma; sarcoma.
OS	Homo sapiens.
PN	WO9931128-A2.
XX	24-JUN-1999.
DD	
PF	02-DEC-1998; 98WO-US25649.
XX	16-DEC-1997; 97US-0991945.
XX	(INCY-) INCYTE PHARM INC.
PA	AU-Young J, Bandman O, Hillman JL, Kaser MR, Tang YT;
PI	WPI; 1999-457916/38.
XX	N-PSDB; X89503.
DR	New tumour necrosis factor-R2-like protein - useful in the treatment
XX	of osteogenesis, developmental, reproductive, immunological and
PT	neoplastic disorders
PT	Claim 1; Fig 1A-C; 81pp; English.
XX	The present sequence represents a human tumour necrosis factor-R2-like
CC	protein (TR2P). The protein is used to treat and prevent osteogenesis,
CC	developmental, reproductive, immunological and neoplastic disorders, and
CC	also to diagnose disorders associated with TR2 protein expression. Such
CC	disorders include osteogenesis disorders such as achondroplasia and
CC	osteoporosis, developmental disorders such as Cushing's syndrome,
CC	muscular dystrophy, and epilepsy, hereditary neuropathies such as
CC	Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism,
CC	hydrocephalus, seizure disorders such as cerebral palsy and spinal
CC	bifida, congenital glaucoma, cataract, or sensorineural hearing loss,
CC	reproductive disorders such as infertility, ovulatory defects and
CC	endometriosis, autoimmune disorders, ectopic pregnancy and teratogenesis,
CC	disruption of spermatogenesis, immunological disorders such as AIDS,
CC	Addison's disease, allergies, bronchitis, atherosclerosis, diabetes
CC	mellitus, Chron's disease, lupus and irritable bowel syndrome, multiple
CC	sclerosis, viral, fungal, helminthic, parasitic and protozoal infections,
CC	and neoplastic disorders including adenocarcinoma, leukaemia, lymphoma,
CC	melanoma, myeloma, sarcoma, and teratocarcinoma.
XX	Sequence 245 AA;
XX	SQ
XX	Query Match 81.2%; Score 667; DB 20; Length 245;
XX	Best Local Similarity 100.0%; Pred. No. 1.5e-51;
XX	Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 AETPTYPWRDAETGERLVCAQCPCGTFVQRCDSPTCGCPPRHYTQFWNYLCRCY 60
DB	
DB	31 AETPTYWRDAETGERLVCAQCPCGTFVQRCDSPTCGCPPRHYTQFWNYLCRCY 90
QY	61 CNVLGGEREEARACHATHNRACRRTGFFAHAGFCLEHASCPPGAGVIAPG 112
DB	
DB	91 CNVLGGEREEARACHATHNRACRRTGFFAHAGFCLEHASCPPGAGVIAPG 142
DE	RESULT 3
XX	ID Y42184 standard; Protein; 271 AA.
XX	AC Y42184;
XX	DT 17-DEC-1999 (first entry)
XX	Human mFLINT #1 protein sequence.
XX	Human; FLINT; mFLINT; OPD; tumour necrosis factor receptor; FasL;
KW	apoptosis; inflammation; cancer; diabetes; acute liver failure;
KW	sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
KW	refrusion-associated injury; aplastic anaemia; differentiation;
KW	growth; myelodysplastic syndrome; pancytopenic condition;
KW	myocardial ischaemia.
OS	Homo sapiens.
PN	WO9950413-A2.
XX	

PD 07-OCT-1999.
 XX 30-MAR-1999; 99WO-US06797.
 XX 30-MAR-1998; 98US-0079856.
 PR 20-MAY-1998; 98US-0086074.
 PR 09-SEP-1998; 98US-0099643.
 PR 17-DEC-1998; 98US-0112577.
 PR 18-DEC-1998; 98US-0112703.
 PR 18-DEC-1998; 98US-0112933.
 PR 22-DEC-1998; 98US-0113407.
 XX (ELIL) LILLY & CO ELI.
 XX Bumol TF, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
 PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
 PI Song HY, Wang J, Wu X, Zuckerman SH;
 XX WPI: 1999-591319/50.
 DR N-PSDB: 225377.
 XX Use of mature FLINT for treating acute liver failure, inflammation,
 PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
 PT and proinflammatory activity
 XX Claim 31; Fig 3; 99pp; English.
 XX The present invention describes therapeutic applications of mature FLINT
 CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
 CC (mFLINT), which is a member of the tumour necrosis factor receptor
 CC superfamily, is used for treating acute liver failure, inflammation of
 CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
 CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
 CC injury or disorder such as hypercoagulation (including use with
 CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
 CC or disorder, Type I diabetes, cancer, cell damage or damage to an
 CC innocent bystander tissue that is induced by a chemotherapeutic agent or
 CC therapeutic irradiation, treating haematopoietic progenitor cells that
 CC have been exposed to therapeutic radiation or chemotherapy, aplastic
 CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
 CC also used for promoting the growth or differentiation of a haematopoietic
 CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
 CC resulting from abnormal myocardial ischaemia. The present sequence
 CC represents human mFLINT.
 XX Sequence 271 AA;
 SQ

Query Match 81.2%; Score 667; DB 20; Length 271;
 Best Local Similarity 100.0%; Pred. No. 1.6e-51;
 Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AETPTYPWRDAETGERLVCAOCPGTFVQRCRRDSTPTCGPCPPRHVYTFQWNYLCRY 60
 Db 2 aetptypwrdaetgerlvcaocpgtgvqrcrrdsptcpcpprhytqfwnylcry 61
 Qy 61 CNVLGGEREEARACHATHNRCACRTGFFAHAGFCLEHASCPCPGAGVIAPG 112
 Db 62 cnvlgereearachathnrcacrtgffahagfclehascpvgviapg 113

RESULT 4
 W66102
 ID W66102 standard; Protein; 300 AA.
 XX
 AC W66102;
 XX
 DT 02-DEC-1998 (first entry)
 XX
 DE Amino acid sequence of tumour necrosis related receptor (TR4).
 XX Human; tumour necrosis related receptor; TR4; agonist; antagonist;
 KW inhibition; chronic; acute; inflammation; arthritis; septicaemia;
 KW

KW autoimmune disease; transplant rejection; stroke; cancer;
 KW Alzheimer's disease.
 OS Homo sapiens.
 XX EP861850-A1.
 XX 02-SEP-1998.
 XX 20-JAN-1998; 98EP-0300382.
 PF 04-FEB-1997; 97US-0794796.
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 PA Emery J, Tan KB, Truneh A, Young PR;
 PI WPI: 1998-508248/44.
 XX N-PSDB: V07654.
 DR New DNA encoding tumour necrosis related receptor - used to treat
 XX and prevent e.g. inflammation, arthritis, septicaemia, autoimmune
 PT diseases, transplant rejection, infection, stroke, ischaemia, ARDS,
 PT restenosis, AIDS, bone disorders and cancer
 XX Claim 1; Fig 1; 2lpp; English.
 XX This is the amino acid sequence of the human tumour necrosis related
 CC receptor (TR4), used in the method of the invention. The TR4 protein
 CC or its agonist can be used to treat a subject in need of enhanced
 CC TR4 polypeptide activity. The antagonist is used to inhibit TR4
 CC polypeptide activity. The active agents can be used for the
 CC treatment and prevention of diseases such as chronic and acute
 CC inflammation, arthritis, septicaemia, autoimmune diseases, transplant
 CC rejection, stroke, cancer, Alzheimer's disease.
 XX Sequence 300 AA;
 SQ

Query Match 81.2%; Score 667; DB 19; Length 300;
 Best Local Similarity 100.0%; Pred. No. 1.8e-51;
 Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AETPTYPWRDAETGERLVCAOCPGTFVQRCRRDSTPTCGPCPPRHVYTFQWNYLCRY 60
 Db 31 aetptypwrdaetgerlvcaocpgtgvqrcrrdsptcpcpprhytqfwnylcry 90
 Qy 61 CNVLGGEREEARACHATHNRCACRTGFFAHAGFCLEHASCPCPGAGVIAPG 112
 Db 91 cnvlgereearachathnrcacrtgffahagfclehascpvgviapg 142

RESULT 5
 W63622
 ID W63622 standard; Protein; 300 AA.
 XX
 AC W63622;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Human tumour necrosis factor receptor-6 alpha protein.
 XX
 KW Human tumour necrosis factor receptor-6 alpha; TNFR-6 beta;
 KW endothelial cells; keratinocytes; normal prostate; apoptosis;
 KW prostate tumour tissue.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..30
 FT Protein 31..300
 FT /note= "TNFR-6 alpha"
 FT Region 31..282

KW reperfusion-associated injury; aplastic anaemia; differentiation;
 KW growth; myelodysplastic syndrome; pancytopenic condition;
 KW myocardial ischaemia.
 XX Homo sapiens.
 OS
 PN WO9950413-A2.
 XX
 XX 07-OCT-1999.
 PD
 XX
 PF 30-MAR-1999; 99WO-US06797.
 XX
 PR 30-MAR-1998; 98US-0079856.
 PR 20-MAY-1998; 98US-0086074.
 PR 09-SEP-1998; 98US-0099643.
 PR 17-DEC-1998; 98US-0112577.
 PR 18-DEC-1998; 98US-0112703.
 PR 18-DEC-1998; 98US-0112933.
 PR 22-DEC-1998; 98US-0113407.
 XX
 PA (ELIL) LILLY & CO ELI.
 XX
 PI Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
 PI Hui KY, Khartonenkov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
 PI Song HY, Wang J, Wu X, Zuckerman SH;
 XX
 XX WPI: 1999-591319/50.
 DR N-PSDB; 225375.
 DR
 XX
 PT Use of mature FLINT for treating acute liver failure, inflammation,
 PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
 PT and proinflammatory activity
 XX
 XX Claim 30; Fig 1; 99pp; English.
 PS
 XX The present invention describes therapeutic applications of mature FLINT
 CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
 CC (mFLINT), which is a member of the tumour necrosis factor receptor
 CC superfamily, is used for treating acute liver failure, inflammation of
 CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
 CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
 CC injury or disorder such as hypercoagulation (including use with
 CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
 CC or disorder, Type I diabetes, cancer, cell damage or damage to an
 CC innocent bystander tissue that is induced by a chemotherapeutic agent or
 CC therapeutic irradiation, treating haematopoietic progenitor cells that
 CC have been exposed to therapeutic radiation or chemotherapy, aplastic
 CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
 CC also used for promoting the growth or differentiation of a haematopoietic
 CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
 CC resulting from abnormal myocardial ischaemia. The present sequence
 CC represents human FLINT.
 XX
 SQ Sequence 300 AA;
 Query Match 81.2%; Score 667; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 1.8e-51;
 Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 AETPTYWRDAETGERLVCAQCPGTFVQRCRRDSPTTCGPPRRHYTQFWNYLCRY 60
 Db 31 aetptywrdaetgerlvcaqcpptgfvrpcrrdspttcgpprrhytqfwnylcr 90
 Qy 61 CNVLGGEREEARACHATHNRCRCRTGFFAHAGFCLEHASCPPGAGVIAPG 112
 Db 91 cnvlcgereearachathnrcrcrtgffahagfclehascpaggviapg 142
 RESULT 8
 Y17479
 ID Y17479 standard; Protein: 300 AA.
 KW

Y17479;
 AC
 XX 02-AUG-1999 (first entry)
 DT
 XX Mammalian tumour necrosis factor receptor OPG-2.
 DE
 XX Tumour necrosis factor receptor; TNF receptor; OPG-2; Paget's disease;
 KW osteopenic disorder; osteoclast activity; primary osteoporosis;
 KW hyperglycaemia; osteolytic metastasis; immune response; cancer.
 XX
 OS Mammalia.
 XX
 XX WO9926977-A1.
 PN
 XX 03-JUN-1999.
 PD
 XX 24-NOV-1998; 98WO-US25065.
 PF 17-FEB-1998; 98US-0074896.
 PR 24-NOV-1997; 97US-0066446.
 PR
 XX (BIOJ) BIOGEN INC.
 PA
 XX
 PI Tschopp J;
 PI
 XX WPI: 1999-347693/29.
 DR N-PSDB; X76052.
 DR
 XX New tumour necrosis factor family receptor OPG-2
 PT
 XX Claim 1; Page 18; 22pp; English.
 PS
 XX The present sequence represents a mammalian tumour necrosis factor
 CC receptor, designated OPG-2. OPG-2, is a member of the tumour necrosis
 CC factor receptor family, and can be used: (i) to raise specific
 CC antibodies (Ab), (ii) to treat osteopenic disorders associated with
 CC excessive osteoclast activity, e.g. primary osteoporosis, Paget's
 CC disease, hyperglycaemia of malignancy, or osteolytic metastases; (iii)
 CC for affinity purification of cognate ligands, and (iv) to screen for
 CC ligands (antagonists or agonists). Ab, or other OPG-2 blocking agents
 CC such as soluble forms of the protein, are used to prevent, or reduce
 CC severity of, an immune response, and for treating cancer. They can also
 CC be used in diagnostic assays. The nucleic acid sequence encoding OPG-2
 CC can be used as a probe to isolate related sequences from other species.
 XX
 SQ Sequence 300 AA;
 Query Match 81.2%; Score 667; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 1.8e-51;
 Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 AETPTYWRDAETGERLVCAQCPGTFVQRCRRDSPTTCGPPRRHYTQFWNYLCRY 60
 Db 31 aetptywrdaetgerlvcaqcpptgfvrpcrrdspttcgpprrhytqfwnylcr 90
 Qy 61 CNVLGGEREEARACHATHNRCRCRTGFFAHAGFCLEHASCPPGAGVIAPG 112
 Db 91 cnvlcgereearachathnrcrcrtgffahagfclehascpaggviapg 142
 RESULT 9
 Y06817
 ID Y06817 standard; Protein: 300 AA.
 XX
 AC Y06817;
 XX
 XX 24-JUN-1999 (first entry)
 DT
 XX Human Dcr3 polypeptide.
 DE
 XX Dcr3 polypeptide; tumour necrosis factor receptor; TNFR; Fas ligand;
 KW apoptosis; T cell mediated immune response; allergy; asthma; cancer;
 KW

RESULT 11	
W95082	W95082 standard; Protein; 300 AA.
XX	
XX	W95082;
XX	
DT	20-MAY-1999 (first entry)
XX	
XX	Orphan receptor (HUMAN NTR-1) polypeptide.
XX	
XX	HUMAN NTR-1; orphan receptor; osteoprotegerin; OPG; TNFR; human;
KW	tumour necrosis factor receptor; muscle disorder; bone mass; screening;
KW	muscle metabolism; binding agent; cognate ligand.
XX	
OS	Homo sapiens.
XX	
PN	W09507738-A2.
XX	
XX	18-FEB-1999.
PD	
XX	04-AUG-1998; 98WO-US16202.
XX	
PR	06-AUG-1997; 97US-0054869.
XX	
PA	(PROC) PROCTER & GAMBLE CO.
PA	(REGE-) REGENERON PHARM INC.
XX	
PI	Masiakowski PJ, Morris J, Valenzuela DM;
XX	
DR	WPI; 1999-167365/14.
DR	N-PSDB; X22300.
XX	
XX	
PT	Novel orphan human receptor polypeptide and nucleic acid - useful as
PT	diagnostic reagents and for treatment of muscle disorders
XX	
XX	Claim 7; Page 21; 23pp; English.
XX	
CC	This represents a HUMAN NTR-1 polypeptide, a novel orphan receptor. The
CC	protein is related to osteoprotegerin (OPG) and to tumour necrosis factor
CC	receptor (TNFR). Host cells transformed with a vector comprising the
CC	HUMAN NTR-1 nucleic acid are used for the recombinant expression of the
CC	protein. HUMAN NTR-1 proteins and antibodies immuno specific for the
CC	protein are useful for diagnosis and treatment of humans and animals,
CC	especially muscle disorders, as the receptor is involved in regulation of
CC	bone mass and muscle metabolism. HUMAN NTR-1 receptors are also useful
CC	for screening for novel binding agents, and cognate ligands, which may be
CC	used to treat disorders associated with HUMAN NTR-1 imbalance.
XX	
SQ	Sequence 300 AA:

Query Match	81.2%	Score 667;	DB 20;	Length 300;
Best Local Similarity	100.0%;	Pred. No. 1.8e-51;		
Matches 112;	Conservative	0;	Mismatches 0;	Gaps 0;
			Indels	0;

Qy	1	AETPTVPRWDAETGERLVCACQPPGFTFVQRP	CRDSTTTCGCP	PRHYTQFWNYL	ERCY	60
31	aetptvprwdaetgerlvcacppqtfvqpcridsp	ttcgcpcpprt	hvtqfwnylery		90	
Db	aetptvprwdaetgerlvcacppqtfvqpcridsp	ttcgcpcpprt	hvtqfwnylery			

Qy 61 CNVLGEREEEARACHATHNRACRCRTGFFAHAGFCLEHASCPGAGVIAPG 112

Db 91 cnvlgereearachathnracrctgffahagfciehascpvgviapg 142

RESULT	12
Y42185	
ID	Y42185 standard; Protein; 273 AA.
XX	
AC	Y42185;
XX	
DT	17-DEC-1999 (first entry)

XX	Human mFLINT #2 protein sequence.
XX	
XX	
XX	Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
KW	apoptosis; inflammation; cancer; diabetes; acute liver failure;
KW	sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
KW	reperfusion-associated injury; aplastic anaemia; differentiation;
KW	growth; myelodysplastic syndrome; pancytopenic condition;
KW	myocardial ischaemia.
XX	
XX	
OS	Homo sapiens.
XX	
XX	WO99504113-A2.
PN	
XX	
XX	07-OCT-1999.
PD	
XX	
XX	30-MAR-1999; 99WO-US06797.
XX	
XX	
XX	30-MAR-1998; 98US-0079856.
PR	
XX	20-MAY-1998; 98US-0086074.
PR	
XX	09-SEP-1998; 98US-0099643.
PR	
XX	17-DEC-1998; 98US-0112577.
PR	
XX	18-DEC-1998; 98US-0112703.
PR	
XX	18-DEC-1998; 98US-0112933.
PR	
XX	22-DEC-1998; 98US-0113407.
PR	
XX	
PA	(ELIL) LILLY & CO ELI.
XX	
XX	Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
PI	Hui KY, Kharitononkov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
PI	Song HY, Wang J, Wu X, Zuckerman SH;
XX	
DR	WPI; 1999-591319/50.
DR	N-PSDB; 725378.
XX	
PT	Use of mature FLINT for treating acute liver failure, inflammation,
PT	cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
PT	and proinflammatory activity
XX	
XX	Example 2; Fig 4; 99pp; English.
PS	
XX	
XX	The present invention describes therapeutic applications of mature FLINT
CC	(mFLINT) for use in the treatment of acute liver failure. Mature FLINT
CC	(mFLINT), which is a member of the tumour necrosis factor receptor
CC	superfamily, is used for treating acute liver failure, inflammation of
CC	the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
CC	with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
CC	injury or disorder such as hypercoagulation (including use with
CC	thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
CC	or disorder, Type I diabetes, cancer, cell damage or damage to an
CC	innocent bystander tissue that is induced by a chemotherapeutic agent or
CC	therapeutic irradiation, treating haematopoietic progenitor cells that
CC	have been exposed to therapeutic radiation or chemotherapy, aplastic
CC	anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
CC	also used for promoting the growth or differentiation of a haematopoietic
CC	progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
CC	resulting from abnormal myocardial ischaemia. The present sequence
CC	represents human mFLINT.
XX	
XX	Sequence 273 AA;
XX	

Query Match	79.9%	Score 656;	DB 20;	Length 273;
Best Local Similarity	98.2%;	Pred. No. 1.5e-50;		
Matches 112:	Conservative	0;	Mismatches	0;
	Indels	2;	Gaps	1;

```

QY 1 AETPTYPWDAETGERLVCACCPGTFVQRCRRDSPITCGPCPRHYTQFWNYLERCRY 60
Db 2 aetptypwdaetgerlvacgcpptgfvrpcrrdspitcgpcprhytqfwnylercry 61
QY 61 CNVLGGEREEARACHATHNRA--CRCRTGFFAHAGFCLHASCPGPGAGVIAPG 112
Db 62 cnvlcgereearachathnrcrcrtgffahagfclhascppgpgagviapg 115

```

Matches 112; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

RESULT 13
Y42183
ID Y42183 standard; Protein: 302 AA.
XX
AC Y42183;
XX
DT 17-DEC-1999 (first entry)
XX
DE Human FLINT #2 protein sequence.
XX
KW Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
KW reperfusion-associated injury; aplastic anaemia; differentiation;
KW growth; myelodysplastic syndrome; pancytopenic condition;
KW myocardial ischaemia.
XX
OS Homo sapiens.
XX
PN W09950413-A2.
XX
PD 07-OCT-1999.
XX
PF 30-MAR-1999; 99WO-US06797.
XX
PR 30-MAR-1998; 98US-0079856.
PR 20-MAY-1998; 98US-0086074.
PR 09-SEP-1998; 98US-0099643.
PR 17-DEC-1998; 98US-0112577.
PR 18-DEC-1998; 98US-0112703.
PR 18-DEC-1998; 98US-0112933.
PR 22-DEC-1998; 98US-0113407.
XX
PA (ELIL) LILLY & CO ELI.
XX
PI Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
PI Song HY, Wang J, Wu X, Zuckerman SH;
XX
WPI: 1999-591319/50.
DR N-PSDB; Z25376.
XX
PT Use of mature FLINT for treating acute liver failure, inflammation,
PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
PT and proinflammatory activity
XX
PS Example 2; Fig 2; 99pp; English.
XX
CC The present invention describes therapeutic applications of mature FLINT
CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
CC (mFLINT), which is a member of the tumour necrosis factor receptor
CC superfamily, is used for treating acute liver failure, inflammation of
CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
CC injury or disorder such as hypercoagulation (including use with
CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
CC or disorder, Type I diabetes, cancer, cell damage or damage to an
CC innocent bystander tissue that is induced by a chemotherapeutic agent or
CC therapeutic irradiation, treating haematopoietic progenitor cells that
CC have been exposed to therapeutic radiation or chemotherapy, aplastic
CC anemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
CC also used for promoting the growth or differentiation of a haematopoietic
CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
CC resulting from abnormal myocardial ischaemia. The present sequence
CC represents human FLINT.
XX
SQ Sequence 302 AA;

Query Match 79.9%; Score 656; DB 20; Length 302;
Best Local Similarity 98.2%; Pred. No. 1.7e-50;

QY 1 AETPTYPWRDAETGERLVCAOCPPGCTFVQRCRDSPTTCGCPRHVYTOFWNYLCRCRY 60
|||||
DB 31 aetptypwrdaetgerlvcaocppgctfvrpcrrdsptcgcprhrytqfwnylcrery 90
|||||
QY 61 CNVLCGEREEERACHATHNRA--CRCRTGFFAHAGFCLHASCPPGAGVIAPG 112
|||||
DB 91 cnvlgereeeearachathnracrcrtgffahagfclhascppgagviapg 144
|||||
RESULT 14
Y77458
ID Y77458 standard; Protein: 300 AA.
XX
AC Y77458;
XX
DT 05-JUN-2000 (first entry)
XX
DE Human TNF receptor-like protein, HDTEA84.
XX
KW TNF receptor family; tumour necrosis factor; HDTEA84; HSLJD37R;
KW Rank-like protein; RANKL; immune disorder; inflammation; allergy;
KW immunosuppressant; antiarthritic; antirheumatoid; antiinflammatory;
KW dermatological; antithyroid.
XX
OS Homo sapiens.
XX
PN W0200001817-A2.
XX
PD 13-JAN-2000.
XX
PF 06-JUL-1999; 99WO-US12366.
XX
PR 06-JUL-1998; 98US-0110938.
PR 13-JUL-1998; 98US-0114466.
PR 23-JUL-1998; 98US-0053897.
PR 12-AUG-1998; 98US-0132968.
PR 18-AUG-1998; 98US-0136214.
PR 11-SEP-1998; 98US-0099999.
XX
PS (SCHE) SCHERING CORP.
XX
PI Bates PEM, Lebecque SJE, Murphy EE, Mattson JD, Gorman DM;
PI Hedrick JA, Wang L, Zlotnik A, Murgolo NJ, Greene JR, Johnston JA;
PI Bazan JF, Mahony D, Lees EM,
XX
WPI: 2000-171015/15.
DR N-PSDB; Z92404.
XX
PT New isolated mammalian genes, used to develop products for treating
PT e.g. immune, inflammatory or allergic abnormalities, cancers or
PT degenerative conditions
XX
PS Claim 24; Page 157; 218pp; English.
XX
CC The invention relates to a number of primate and/or rodent proteins, and
CC the genes which encode them. The invention encompasses human dendritic
CC cell prostaglandin transporter (DC-PGT); the TNF (tumour necrosis
CC factor) receptor family-related proteins HDTEA84, HSLJD37R and RANKL;
CC human CC chemokine HCC5; human deubiquitinating proteins Dub11 and Dub
CC 12; human MD-1 and human and murine MD-2 proteins, which exhibit the
CC properties of ligands for proteins comprising a leucine-rich motif (LRK);
CC human cyclin E2; cDNAs encoding these proteins; and antibodies against
CC these proteins. The proteins can be used for modulating the physiology or
CC development of a cell. They can be used to mediate uptake of substrates
CC (e.g., prostaglandin-like molecules), to modulate or mediate cellular
CC interactions (e.g., induce or prevent trafficking, proliferation, or
CC differentiation of cells), or are intracellular proteins which are
CC important in various cellular processes such as the deubiquitination of
CC proteins or cell cycle regulation. The products can be used for
CC treating medical conditions such as immune, inflammatory or allergic
CC disorders, or abnormal cellular proliferation, for example, cancers or

A; Cross-references: GB:M60469; NID:g199827; PIDN:AAA39752.1; PID:g199828
R:Goodwin, R.G.; Anderson, D.; Jerzy, R.; Davis, T.; Brannan, C.I.; Cope
Mol. Cell. Biol. 11, 3020-3026, 1991

```
RESULT 4
D72175
G2R protein - variola minor virus (strain Garcia-1966)
C:Species: variola minor virus
C:Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 20-Jun-2000
C:Accession: D72175
R:Shchelkunov, S.N.; Totmenin, A.V.; Gutorov, V.V.; Safronov, P.F.; Massung, R.F.; LO
submitted to GenBank, March 1998
A:Description: Analysis of the complete coding sequence of DNA of alastrim variola m
A:Reference number: A72150
A:Accession: D72175
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-349 <SHC>
A:Cross-references: GB:V16780; NID:g5830555; PIDN:CAB54798.1; PID:g5830759
A:Experimental source: strain Garcia-1966
C:Genetics:
A:Gene: G2R
C:Superfamily: myxoma virus T2 protein; NGF receptor repeat homology

Query Match 23.3% Score 191.5; DB 2; Length 349;
Best Local Similarity 33.9% Pred. No. 4.le-09;
Matches 38; Conservative 12; Mismatches 55; Indels 7; Gaps 2;

QY 4 PTYPWRDAE-TGERLVCAQCPCPGTFVQRCRRDSTTCGCPGPPRHYTQFWNYLERCRYCN 62
DB 28 PNGCKDTEYKRHNLCCLSCPPTGYASRLCDSTNTQCTPCGSGTFTSRNNHLPACLSN 87
QY 63 VLCGEREEARACHATHNRACRGTGFF-----AHAGFCLEHASCPCPPGAGV 108
DB 88 GRCNSNQVETRSCNTHNRICECPGYICLLKSGSGCKACVSTKCGIGYGV 139

RESULT 5
D36858
gene G4R protein - variola virus
N:Alternate names: B28R protein (COP)
C:Species: variola virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 08-Oct-1999
C:Accession: D36858; S46888; S32385; S35987
R:Blinov, V.M.
submitted to GenBank, November 1992
A:Description: not shown.
A:Reference number: A36859
A:Accession: D36858
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-349 <BLI>
A:Cross-references: GB:X69198; NID:g456758; PIDN:CAA49137.1; PID:g457087
A:Experimental source: strain India-1967, ssp. major, isolate Ind3
R:Kolykhalov, A.A.; Blinov, V.M.; Gytarov, V.V.; Pozdnyakov, S.G.; Chizhikov, V.E.; F
submitted to the EMBL data Library, April 1992
A:Description: Nucleotide sequence analysis of the region of variola virus XhoI F O H
A:Reference number: S46868
A:Accession: S46888
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-349 <ROL>
A:Cross-references: EMBL:X67117; NID:g516428; PIDN:CAA47540.1; PID:g516449
A:Experimental source: strain India-1967, isolate Ind3
R:Shchelkunov, S.N.; Blinov, V.M.; Sandakhchiev, L.S.
FEBS Lett. 319, 80-83, 1993
A:Title: Genes of variola and vaccinia viruses necessary to overcome the host protect
A:Reference number: S32385; MUID:93202281
A:Accession: S32385
A:Molecule type: DNA
A:Residues: 31-168 <SHC>
A:Cross-references: EMBL:X69198
A:Experimental source: strain India-1967, ssp. major
C:Genetics:
A:Gene: G4R
C:Superfamily: myxoma virus T2 protein; NGF receptor repeat homology
```

```
RESULT 4
D72175
G2R protein - variola minor virus (strain Garcia-1966)
C:Species: variola minor virus
C:Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 20-Jun-2000
C:Accession: D72175
R:Shchelkunov, S.N.; Totmenin, A.V.; Gutorov, V.V.; Safronov, P.F.; Massung, R.F.; LO
submitted to GenBank, March 1998
A:Description: Analysis of the complete coding sequence of DNA of alastrim variola m
A:Reference number: A72150
A:Accession: D72175
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-349 <SHC>
A:Cross-references: GB:V16780; NID:g5830555; PIDN:CAB54798.1; PID:g5830759
A:Experimental source: strain Garcia-1966
C:Genetics:
A:Gene: G2R
C:Superfamily: myxoma virus T2 protein; NGF receptor repeat homology

Query Match 23.3% Score 191.5; DB 2; Length 349;
Best Local Similarity 33.9% Pred. No. 4.le-09;
Matches 38; Conservative 12; Mismatches 55; Indels 7; Gaps 2;

QY 4 PTYPWRDAE-TGERLVCAQCPCPGTFVQRCRRDSTTCGCPGPPRHYTQFWNYLERCRYCN 62
DB 28 PNGCKDTEYKRHNLCCLSCPPTGYASRLCDSTNTQCTPCGSGTFTSRNNHLPACLSN 87
QY 63 VLCGEREEARACHATHNRACRGTGFF-----AHAGFCLEHASCPCPPGAGV 108
DB 88 GRCNSNQVETRSCNTHNRICECPGYICLLKSGSGCKACVSTKCGIGYGV 139

RESULT 5
D36858
gene G4R protein - variola virus
N:Alternate names: B28R protein (COP)
C:Species: variola virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 08-Oct-1999
C:Accession: D36858; S46888; S32385; S35987
R:Blinov, V.M.
submitted to GenBank, November 1992
A:Description: not shown.
A:Reference number: A36859
A:Accession: D36858
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-349 <BLI>
A:Cross-references: GB:X69198; NID:g456758; PIDN:CAA49137.1; PID:g457087
A:Experimental source: strain India-1967, ssp. major, isolate Ind3
R:Kolykhalov, A.A.; Blinov, V.M.; Gytarov, V.V.; Pozdnyakov, S.G.; Chizhikov, V.E.; F
submitted to the EMBL data Library, April 1992
A:Description: Nucleotide sequence analysis of the region of variola virus XhoI F O H
A:Reference number: S46868
A:Accession: S46888
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-349 <ROL>
A:Cross-references: EMBL:X67117; NID:g516428; PIDN:CAA47540.1; PID:g516449
A:Experimental source: strain India-1967, isolate Ind3
R:Shchelkunov, S.N.; Blinov, V.M.; Sandakhchiev, L.S.
FEBS Lett. 319, 80-83, 1993
A:Title: Genes of variola and vaccinia viruses necessary to overcome the host protect
A:Reference number: S32385; MUID:93202281
A:Accession: S32385
A:Molecule type: DNA
A:Residues: 31-168 <SHC>
A:Cross-references: EMBL:X69198
A:Experimental source: strain India-1967, ssp. major
C:Genetics:
A:Gene: G4R
C:Superfamily: myxoma virus T2 protein; NGF receptor repeat homology
```

F:32-66/Domain: NGF receptor repeat homology <NGF>
F:68-109/Domain: NGF receptor repeat homology <NG2>
F:110-151/Domain: NGF receptor repeat homology <NG3>

Query Match	23.3%	Score 191.5;	DB 2;	Length 349;
Best Local Similarity	33.9%	Pred. No. 4.1e-09;		
Matches	38;	Conservative	12;	Mismatches 55;
				Indels 7;
				Gaps 2;

Qy 4 PTYPWRDAE-TGERLVCAQCPCGTFTVQRPCRDRSDPTTCGCPPRHYTQFWNYLERCRYCN 62
- : | - | | | | | : | | | | : | | | | : | | | |
Dd 28 PNGCKDKTEYKRHNLCCLSCPGETYASRLCDSKTNQTCTCGSGTFTSRNNHLPACLSCN 87

Qy 63 VLGEREEERACHATHNNRACRTGFF-----AHAGFCLEHASCPPGAGV 108
| : | | : | | | | | : : :
Db 88 GRCNSNOVETPSCNTTHNRICECSPGYCLLKSSGGCKACVSOTKCGIGYGV 139

RESULT 6
A35356
tumor necrosis factor receptor type 2 precursor - human
N:Alternate names: 75K tumor necrosis factor receptor
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A35356; A36475; A48416; A36007; A23666; B35010; I38094
R:Smith, C.A.; Davis, T.; Anderson, D.; Solam, L.; Beckmann, M.P.; Jerzy, R.; Dower, S.K.
Science 248, 1019-1023, 1990
A:Title: A receptor for tumor necrosis factor defines an unusual family of cellular and
A:Reference number: A35356; MUID:90260639
A:Accession: A35356
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-461 <SMI>
A:Cross-references: GB:M32315; NID:g189185; PIDN:AAA59929.1; PID:g189186
R:Kohn, T.; Brewer, M.T.; Baker, S.L.; Schwartz, P.E.; King, M.W.; Hale, K.K.; Squires,
Proc. Natl. Acad. Sci. U.S.A. 87, 8331-8335, 1990
A:Title: A second tumor necrosis factor receptor gene product can shed a naturally occurring
A:Reference number: A36475; MUID:91045991
A:Accession: A36475
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-195, 'R', 197-461 <KOH>
A:Cross-references: GB:M55994; GB:M38549; NID:g339757; PIDN:AAA36755.1; PID:g339758
R:Dambic, Z.; Loetscher, H.; Gubler, U.; Pan, Y.C.; Lahm, H.W.; Gentz, R.; Brockhaus, M.
Cytokine 2, 231-237, 1990
A:Title: Two human TNF receptors have similar extracellular, but distinct intracellular,
A:Reference number: A48416; MUID:91370690

A:Reference number: A35010; MUID:90110215
A:Accession: B35010
A:Status: preliminary
A:Molecule type: protein
A:Residues: 27-31 <ENG>
R:Kuhnert, P.; Kemper, O.; Wallach, D.
Gene 150, 381-386, 1994
A:Title: Cloning, sequencing and partial functional characterization of the 5' region
A:Reference number: I38094; MUID:95121934
A:Accession: I38094
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-37 <RES>
A:Cross-references: EMBL:X80021; NID:g6656044; PIDN:CAA56324.1; PID:g825701
C:Genetics:
A:Gene: GDB:TNFR2
A:Cross-references: GDB:125914; OMIM:191191
A:Map position: 1p36.2-1p36.2
A:Introns: 26/3
A:Note: the list of introns is incomplete
C:Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology
C:Keywords: duplication; glycoprotein; receptor; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-416/Product: tumor necrosis factor receptor type 2 #status experimental <MAT>
F:40-76/Domain: NGF receptor repeat homology <NG1>
F:78-119/Domain: NGF receptor repeat homology <NG2>
F:120-162/Domain: NGF receptor repeat homology <NG3>
F:164-201/Domain: NGF receptor repeat homology <NG4>
F:262-279/Domain: transmembrane #status predicted <TMN>
F:280-461/Domain: intracellular #status predicted <INT>
F:171-193/Binding site: carbohydrate (Asn) (covalent) #status predicted

```

Query Match      23.0%; Score 189; DB 1; Length 461;
Best Local Similarity 33.0%; Pred. No. 8.2e-09;
Matches 34; Conservative 12; Mismatches 51; Indels 6; Gaps 1;

Qy 16 RLVCAQCPPTTFVORPCRDRSPPTTCGCPPPHYTFQFNWYLCRYCNVLCGEREEARAC 75
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 51 QMCCSCSPGQAHAFVCTKTSDTVCDCEDSTYTLQLMNVPECLSCGSRCSQDVETQAC 110
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 76 HATINRACRRTGFFAHAG-----FCLHACSPPGAGVIAPG 112
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Db 111 TREONR1CTCRPGWYALSKEGCRCLCAPLKRKCRPGFVARPG 153

```

```

RESULT      7
I541182
tumor necrosis factor receptor 2-related protein - human
C:Species: Homo sapiens (man)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 17-Mar-2000
C:Accession: I541182
R:Baens, M.; Chaifanet, M.; Cassiman, J.J.; Van den Berghe, H.; Marynen, P.
Genomics 16, 214-218, 1993
A:Title: Construction and evaluation of a hncDNA library of human l2p transcribed seq
A:Reference number: I541182; MUID:93252381
A:Accession: I541182
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-435 <RES>
A:Cross-references: GB:L04270; NID:g339761; PIDN:AAA36757.1; PID:g339762
C:Genetics:
A:Gene: GDB:LTBR
A:Cross-references: GDB:1230195; OMIM:600979
A:Map position: 12p13.3-12p13.1

```

Query Match 22.7% Score 186; DB 2; Length 435;
Best Local Similarity 37.9% Pred. No. 1.4e-08;
Matches 39; Conservative 12; Mismatches 40; Indels 12; Gaps 5;
QV 12 ETGERLVCAOCPGTFVORPCRDSPPTTCGCPRHVYTFWNVL---BRCRYCNVLCGER 68

Db 52 EPQHRCSCRPPTGVYVSAKSRIRDTVCATCAENSYNEHWNLYTICQLCRPCDPVVG-- 109
QY 69 EEARACHATHNRACRCRTGFFAHAGFLE--H---ASCPPG 105
Db 110 LEEIAPCTSKRKTQCRQCPGMFC-AAWALECTHCELLSDCPG 151

RESULT 8
B43692
T2 protein - rabbit fibroma virus
C:Species: rabbit fibroma virus, Shope fibroma virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: B43692
R:Upton, C.; DeLange, A.M.; McFadden, G.
Virolgy 160, 20-30, 1987
A:Title: Tumorigenic poxviruses: genomic organization and DNA sequence of the telomeric
A:Reference number: A43692; MUID:87321103
A:Accession: B43692
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-325 <UPT>
A:Cross-references: GB:M17433
C:Superfamily: myxoma virus T2 protein; NGF receptor repeat homology
F:64-105/Domain: NGF receptor repeat homology <NG2>
F:106-147/Domain: NGF receptor repeat homology <NG3>

Query Match 20.8%; Score 170.5; DB 2; Length 325;
Best Local Similarity 31.4%; Pred. No. 2.2e-07;
Matches 37; Conservative 8; Mismatches 58; Indels 15; Gaps 2;

QY 17 LVCAQCPPTGVQRCRRDSTTCGCPRRHYTFQFNYLRCRYCNVLCGEREEARACH 76
Db 38 LCCASCHPGFVASRLCGPGSTVTCSPCEDGTFSTAHAPACVSCRGPTGHLSESQPCD 97
QY 77 ATHNRACRCRTGFFA-----HAGFLEHASCPPGAGVIAPGESWARGGAPRSGRR 128
Db 98 RTHDRVCNCSGTGNYCLLKQNGCRICAPQTKCPAGYGV-----SGHTRAGDTLC 146

RESULT 9
GOVZML
T2 protein - myxoma virus (strain Lausanne)
C:Species: myxoma virus
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 18-Jun-1999
C:Accession: A40566
R:Upton, C.; Macen, J.L.; Schreiber, M.; McFadden, G.
Virolgy 184, 370-382, 1991
A:Title: Myxoma virus expresses a secreted protein with homology to the tumor necrosis
A:Reference number: A40566; MUID:91335768
A:Accession: A40566
A:Molecule type: DNA
A:Residues: 1-326 <UPT>
A:Cross-references: GB:M95181; GB:M37976; NID:g332309; PID:AAA46632.1; PID:g332310
C:Superfamily: myxoma virus T2 protein; NGF receptor repeat homology
C:Keywords: glycoprotein
F:64-105/Domain: NGF receptor repeat homology <NG2>
F:106-147/Domain: NGF receptor repeat homology <NG3>
F:66,181,205,238/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 20.2%; Score 165.5; DB 1; Length 326;
Best Local Similarity 29.7%; Pred. No. 5.8e-07;
Matches 35; Conservative 10; Mismatches 58; Indels 15; Gaps 2;

QY 17 LVCAQCPPTGVQRCRRDSTTCGCPRRHYTFQFNYLRCRYCNVLCGEREEARACH 76
Db 38 LCCTSCPPGSYASRLCGPGSDTVCSPCKNETFTASTNHAPACVSCRGRTGHLSESQSD 97
QY 77 ATHNRACRCRTGFFA-----HAGFLEHASCPPGAGVIAPGESWARGGAPRSGRR 128
Db 98 KTRDRVCDCSAGNYCLLKQGGCRICAPKTKCPAGYGV-----SGHTRTGDTLC 146

RESULT 10
I37552
OX40 homolog - human
C:Species: Homo sapiens (man)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 11-Jan-2000
C:Accession: I37552
R:Latza, U.; Durkop, H.; Schnittger, S.; Ringeling, J.; Eitelbach, F.; Hummel, M.; Fo
Eur. J. Immunol. 24, 677-683, 1994
A:Title: The human OX40 homolog: CDNA structure, expression and chromosomal assignmen
A:Reference number: I37552; MUID:94170844
A:Accession: I37552
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-277 <RES>
A:Cross-references: EMBL:X75962; NID:g472957; PIDN:CAA53576.1; PID:g472958
C:Superfamily: CD27 antigen; NGF receptor repeat homology

Query Match 17.9%; Score 147; DB 2; Length 277;
Best Local Similarity 31.9%; Pred. No. 1.8e-05;
Matches 37; Conservative 14; Mismatches 51; Indels 14; Gaps 5;

QY 5 TYPWRDAFTGERLVCAQCPPTGVQRCRRDSTTCGCPRRHYTFQFNY--LERCRYCN 62
Db 35 TYPNSNDR-----CCECRPGNGMYVSCRSQNTVCRPGFGFYNDVVSSKPKPCPTWCN 88
QY 63 VLCGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPPGAGVIAPGESWA 116
Db 89 LRSG--SERKQCLTATQTVTCRCRAGTQPLDSYKPGVDCAPCPGH--FSPGDNOA 140

RESULT 11
A46517
CD27 antigen precursor - human
N:Alternate names: CD27L receptor; T cell activation antigen CD27
C:Species: Homo sapiens (man)
C:Date: 18-Jun-1993 #sequence_revision 22-Apr-1995 #text_change 22-Jun-1999
C:Accession: A46517; A46454
R:Loenen, W.A.; Gravestien, L.A.; Beumer, S.; Mellef, C.J.; Hagemeijer, A.; Borst, J.
J. Immunol. 149, 3937-3943, 1992
A:Title: Genomic organization and chromosomal localization of the human CD27 gene.
A:Reference number: A46517; MUID:93094588
A:Accession: A46517
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-260 <LOE>
A:Note: sequence extracted from NCBI backbone (NCBIP:120386)
A:Note: authors propose an alternative repeat pattern
R:Camerini, D.; Walz, G.; Loenen, W.A.; Borst, J.; Seed, B.
J. Immunol. 147, 3165-3169, 1991
A:Title: The T cell activation antigen CD27 is a member of the nerve growth factor/tu
A:Reference number: A46454; MUID:92013149
A:Accession: A46454
A:Molecule type: mRNA
A:Residues: 1-58, 'A' 60-260 <CAM>
A:Cross-references: GB:M63928; NID:gl80084; PIDN:AAA58411.1; PID:g180085
A:Note: sequence extracted from NCBI backbone (NCBIN:60285, NCBIP:60289)
C:Comment: A soluble CD27 found in serum and urine is formed by proteolysis.
C:Genetics:

A:Gene: GDB:CD27
A:Cross-references: GDB:132582; OMIM:186711
A:Map position: 12p13-12p13
A:Introns: 46/1; 90/1; 150/1; 180/1; 220/1
C:Superfamily: CD27 antigen; NGF receptor repeat homology
C:Keywords: duplication; glycoprotein; homodimer; phosphoprotein; receptor; surface a
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-260/Product: CD27 antigen #status predicted <MAT>
F:21-191/Domain: extracellular #status predicted <EXT>
F:27-63/Domain: NGF receptor repeat homology <NG1>
F:65-103/Domain: NGF receptor repeat homology <NG2>
F:121-188/Region: proline/serine/threonine-rich

Ed. J. Immunol. 23, 920-930, 1995
A>Title: Gene structure and chromosomal localization of the mouse homologue of rat OX40
A;Reference number: I48334;MUID:95255413

QY 92 HAGFCLEHASC-----PPGAGVIAPGES 114
 Db 175 RESECVPCSHCKKNECKMLCLPPPLANVTNPODS 209

Search completed: January 30, 2001, 16:50:33
 Job time: 723 sec

tumor necrosis factor receptor 1 precursor - mouse
 N:Alternate names: tumor necrosis factor receptor, 55K
 C:Species: Mus musculus (house mouse)
 C>Date: 30-Jun-1992 #sequence revision 30-Jun-1992 #text_change 22-Jun-1999
 C:Accession: A38634; B40254; S16677; S19021; I54532
 R:Lewis, M.; Tartaglia, L.A.; Lee, A.; Bennett, G.L.; Rice, G.C.; Wong, G.H.W.; Chen, E.
 Proc. Natl. Acad. Sci. U.S.A. 88, 2830-2834, 1991
 A:Title: Cloning and expression of cDNAs for two distinct murine tumor necrosis factor
 A:Reference number: A38634; MUID:91187885
 A:Accession: A38634
 A:Molecule type: mRNA
 A:Residues: 1-454 <LEW>
 A:Cross-references: GB:M60468; NID:g199825; PIDN:AAA39751.1; PID:g199826
 R:Goodwin, R.G.; Anderson, D.; Jerzy, R.; Brannan, C.I.; Copeland, N.G.; Jenk
 Mol. Cell. Biol. 11, 3020-3026, 1991
 A:Title: Molecular cloning and expression of the type 1 and type 2 murine receptors for
 A:Reference number: A40254; MUID:91246168
 A:Accession: B40254
 A:Molecule type: mRNA
 A:Residues: 1-454 <GO2>
 A:Cross-references: GB:M60468; NID:g199825; PIDN:AAA39751.1; PID:g199826
 R:Barrett, K.; Taylor-Fishwick, D.A.; Cope, A.P.; Kissnerghis, A.M.; Gray, P.W.; Feldma
 Eur. J. Immunol. 21, 1649-1656, 1991
 A:Title: Cloning, expression and cross-linking analysis of the murine p55 tumor necrosis
 A:Reference number: S16677; MUID:91285014
 A:Accession: S16677
 A:Molecule type: mRNA
 A:Residues: 1-454 <BAR>
 A:Cross-references: EMBL:X59238; NID:g53578; PIDN:CAA41922.1; PID:g53579
 R:Rothe, J.G.; Brockhaus, M.; Gentz, R.; Lesslauer, W.
 Immunogenetics 34, 338-340, 1991
 A:Title: Molecular cloning and expression of the mouse Tnf receptor type b.
 A:Reference number: S19021; MUID:92039815
 A:Accession: S19021
 A:Molecule type: mRNA
 A:Residues: 1-454 <ROT>
 A:Cross-references: EMBL:X57796; NID:g54848; PIDN:CAA40936.1; PID:g54849
 R:Bebo, B.F.
 Immunogenetics 39, 450-451, 1994
 A:Title: Nucleotide sequence of the TNF type I receptor from a mouse endothelioma cell
 A:Reference number: I54532; MUID:94245292
 A:Accession: I54532
 A:Status: translated from GB/EMBL/DBDJ
 A:Molecule type: mRNA
 A:Residues: 1-454 <RES>
 A:Cross-references: GB:L26349; NID:g430732; PIDN:AAA59361.1; PID:g430733
 C:Comment: This protein is one of two distantly related receptors for both TNF-alpha (ca
 C:Superfamily: tumor necrosis factor receptor type 1; NGF receptor repeat homology
 C:Keywords: duplication; glycoprotein; receptor; transmembrane protein
 F:1-29/Domain: signal sequence #status predicted <SIG>
 F:30-454/Product: tumor necrosis factor receptor type 1 #status predicted <MAT>
 F:30-212/Domain: extracellular #status predicted <EXT>
 F:44-82/Domain: NGF receptor repeat homology <NG1>
 F:84-126/Domain: NGF receptor repeat homology <NG2>
 F:127-167/Domain: NGF receptor repeat homology <NG3>
 F:168-204/Domain: NGF receptor repeat homology <NG4>
 F:213-235/Domain: transmembrane #status predicted <MEM>
 F:236-454/Domain: intracellular #status predicted <INT>
 F:34,151,202/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 15.3%; Score 125.5; DB 1; Length 454;
 Best Local Similarity 22.6%; Pred. NO. 0.0016;
 Matches 35; Conservative 14; Mismatches 47; Indels 59; Gaps 5;

QY 17 LVCAQCPPTFFVPCRRDSP--TTCGPPPHHYTFWNYLRCRYC----- 61
 Db 57 ICCTCKHKGYLVSDC--PSPGRDVCRECEKGTFTASQNYLRQCLSCCKTKCKEMSQVEI 114
 QY 62 -----NVLCGRREEAR-----ACHATHNRACRCRTGFFA 91
 Db 115 SPCQADKDTVCCKCKENQFQYLSLSETHFCQVDCSPCFNGTGTIPCKETQNTQVNCNAGFFL 174

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 17:06:07 ; Search time 80.04 seconds
(without alignments)
56.486 Million cell updates/sec

Title: US-09-518-931-4_COPY_31_L70

Perfect score: 821

Sequence: 1 AETPTYPWRDAETGERLVCA.....PRSGRRGRCGQVAGPSLAP 140

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	203	24.7	474	1	TNR2_MOUSE
2	191.5	23.3	349	1	VC22_VARV
3	189	23.0	461	1	TNR2_HUMAN
4	186	22.7	435	1	TNR2_HUMAN
5	170.5	20.8	325	1	VT2_SEVKA
6	165.5	20.2	326	1	VT2_MXVVL
7	157	19.1	415	1	TNR2_MOUSE
8	147	17.9	277	1	OX40_HUMAN
9	133	16.2	332	1	FASA_PIG
10	129	15.7	271	1	OX40_RAT
11	128	15.6	272	1	OX40_MOUSE
12	127	15.5	323	1	FASA_BOVIN
13	127	15.5	327	1	FASA_MOUSE
14	126.5	15.4	260	1	CD27_HUMAN
15	125.5	15.3	454	1	TNR1_MOUSE
16	123.5	15.0	595	1	CD30_HUMAN
17	122	14.9	256	1	41BB_MOUSE
18	122	14.9	335	1	FASA_HUMAN
19	120.5	14.7	471	1	TNR1_BOVIN
20	120	14.6	277	1	CD40_HUMAN
21	119	14.5	324	1	FASA_RAT
22	118.5	14.4	255	1	41BB_HUMAN
23	117.5	14.3	250	1	CD27_MOUSE
24	117	14.3	269	1	CD40_BOVIN
25	114.5	13.9	461	1	TNR1_RAT
26	111.5	13.6	416	1	NGFR_CHICK
27	110	13.4	289	1	CD40_MOUSE
28	108.5	13.2	455	1	TNR1_HUMAN
29	106.5	13.0	425	1	NGFR_RAT
30	101.5	12.4	461	1	TNR1_PIG
31	99.5	12.1	427	1	NGFR_HUMAN
32	98	11.9	727	1	NETA_DROME
33	94	11.4	915	1	PAC6_RAT

ALIGNMENTS

RESULT 1

TNR2_MOUSE

ID TNR2_MOUSE STANDARD; PRT; 474 AA.

AC P25119; P97893; 11.1 1964 1 NTC4_MOUSE

DT 01-MAY-1992 (Rel. 22, Created) 1798 1 LMB2_HUMAN

DT 01-MAY-1992 (Rel. 22, Last sequence update) 417 1 WSL1_HUMAN

DT 15-JUL-1999 (Rel. 38, Last annotation update) 915 1 PAC6_MOUSE

DE TUMOR NECROSIS FACTOR RECEPTOR 2 PRECURSOR (TNF-R2) (P75). 3707 1 PGBM_MOUSE

GN TNFRSF1B OR TNFR2 OR TNFR-2. 38 1 TSP2_HUMAN

OS Mus musculus (Mouse). 1172 1 TSP1_XENLA

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; 1173 1 ZAN_PIG

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 10.5 103 1 VA53_VACCC

OC Mus musculus (Mouse). 10.5 103 1 VA53_VACCV

OC Mus musculus (Mouse). 10.5 103 1 LMG2_MOUSE

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

OC Mus musculus (Mouse). 10.5 1801 1 LMB2_RAT

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed, usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: X69198; CAA49137.1; -;
CC EMBL: X67117; CAA47540.1; -;
CC PIR: D36858; D36858.
CC PIR: S35987; S35987.
CC PIR: S46888; S46888.
CC HSP: P19438; INCF.
CC INTERPRO: IPR001368; -;
CC PFAM: PF00020; TNFR_C6; 2.
CC PROSITE: PS00652; TNFR_NGFR_1; 2.
CC PROSITE: PS00050; TNFR_NGFR_2; 2.
CC Repeat.
KW DOMAIN 31 108 2 X TNFR-CYS.
FT REPEAT 31 66 TNFR-CYS 1.
FT REPEAT 67 108 TNFR-CYS 2.
SQ SEQUENCE 349 AA; 38189 MW; D45D0B5C6E780EF CRC64;

Query Match 23.3%; Score 191.5; DB 1; Length 349;
Best Local Similarity 33.9%; Pred. No. 6.5e-10;
Matches 38; Conservative 12; Mismatches 55; Indels 7; Gaps 2;

QY 4 PTYWRDAE-TGERLVCAQCPPTGVQPCRRDSTTCGCPGPRHYTQFWNYLERCRYCN 62
DB 28 PNGCKCKTFYRHNLCCLSCPPGYASRLCDSKTNTCTCGSGTFTSRNHLPACLSCN 87
QY 63 VLGEREEEARACHATHNRACRTGTF-----AHAGFCLERHSCPPGAGV 108
DB 88 GRCSNQVETRSCNTHNRICEGSPGYVLLKSGSGCKACVSTQKCGIGYGV 139

RESULT 3
ID TNR2_HUMAN STANDARD; PRT; 461 AA.
AC P20333;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR 2 PRECURSOR (TUMOR NECROSIS FACTOR
DE BINDING PROTEIN 2) (TBPII) (P80) (TNF-R2) (P75) (CD120B) (ETANERCEPT).
GN TNFRSF1B OR TNFR2 OR TNFR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-90260639; PubMed-2160731;
RX Smith C.A., Davis T., Anderson D., Solam L., Beckmann M.P., Jerzy R.,
RA Dower S.K., Cosman D., Goodwin R.G.;
RT "A receptor for tumor necrosis factor defines an unusual family of
RT cellular and viral proteins.";
RL Science 248:1019-1023(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-91045991; PubMed-2172983;
RX Kohno T., Brewer M.T., Baker S.L., Schwartz P.E., King M.W.,
RA Hale K.K., Squires C.H., Thompson R.C., Vannice J.L.;
RT "A second tumor necrosis factor receptor gene product can shed a
RT naturally occurring tumor necrosis factor inhibitor.";
RT Proc. Natl. Acad. Sci. U.S.A. 87:8331-8335(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE-96299745; PubMed-8661109;
RX Beltinger C.P., White P.S., Maris J.M., Sulman E.P., Jensen S.J.,
RA Lepaslier D., Stallard B.J., Goeddel D.V., Desauvage F.J.,
RA Brodeur G.M.;
RT "Physical mapping and genomic structure of the human TNFR2 gene.";
RT Genomics 35:94-100(1996).
RL

HSP: P19438; INCF.
MGD: MGI:1314883; TNFRSF1B.
INTERPRO: IPR001368; -;
PFAM: PF00020; TNFR_C6; 4.
PROSITE: PS00652; TNFR_NGFR_1; 2.
PROSITE: PS00050; TNFR_NGFR_2; 3.
KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 22
FT CHAIN 23 474 TUMOR NECROSIS FACTOR RECEPTOR 2.
FT DOMAIN 23 258 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 259 288 POTENTIAL.
FT CYTOPLASMIC (POTENTIAL).
FT TNFR-CYS 39 203 4 X TNFR-CYS.
FT REPEAT 39 77 TNFR-CYS 1.
FT REPEAT 78 119 TNFR-CYS 2.
FT REPEAT 120 164 TNFR-CYS 3.
FT REPEAT 165 203 TNFR-CYS 4.
FT DISULFID 40 54 BY SIMILARITY.
FT DISULFID 55 68 BY SIMILARITY.
FT DISULFID 58 76 BY SIMILARITY.
FT DISULFID 79 94 BY SIMILARITY.
FT DISULFID 97 111 BY SIMILARITY.
FT DISULFID 101 119 BY SIMILARITY.
FT DISULFID 121 127 BY SIMILARITY.
FT DISULFID 136 145 BY SIMILARITY.
FT DISULFID 139 163 BY SIMILARITY.
FT DISULFID 166 181 BY SIMILARITY.
FT CARBOHYD 69 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 195 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 474 AA; 50319 MW; 462AE398CAD6563 CRC64;

Query Match 24.7%; Score 203; DB 1; Length 474;
Best Local Similarity 31.8%; Pred. No. 8.6e-11;
Matches 42; Conservative 13; Mismatches 55; Indels 22; Gaps 3;

QY 16 RLVCAQCPPTGVQPCRRDSTTCGCPGPRHYTQFWNYLERCRYCNVLGGEREEARAC 75
DB 52 QMCCAKCPGQVYKFNCKTSDTCADCEASMYTOVWQFRTCLSCSSCTTDOVEIRAC 111
QY 76 HATHNRACRTGTF-----AHAGF---CLEHACGPPGAGVIAPGESWARGAPRSGGRR 128
DB 112 TQQNRVCAACEAGRYCALKTHSGSCRCMRLSKCGPGFGV-----ASSRA 156
QY 129 GRGQVAGPSLAP 140
DB 157 PNGNVLCACAP 168
RESULT 2
ID VC22_VARV STANDARD; PRT; 349 AA.
AC P34015;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE PROTEIN C22/B28 HOMOLOG.
GN Variola virus.
OS Variola virus.
OC dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=INDIA-1967 / ISOLATE IND3;
RX MEDLINE-93202281; PubMed-8384129;
RX Shchelkunov S.N., Blinov V.M., Sandakhchiev L.S.;
RT "Genes of variola and vaccinia viruses necessary to overcome the host
RT protective mechanisms.";
RL FEBS Lett. 319:80-83(1993).
CC -1- SIMILARITY: CONTAINS 2 LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGIONS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

[4] SEQUENCE OF 116-461 FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=90349572; PubMed=2166946;
RA Heller R.A., Song K., Onasch M.A., Fischer W.H., Chang D.,
RA Ringold G.M.;
RT "Complementary DNA cloning of a receptor for tumor necrosis factor
RT and demonstration of a shed form of the receptor.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:6151-6155(1990).
[5]
RX SEQUENCE OF 27-31.
RX MEDLINE=90110215; PubMed=2153136;
RA Engelmann H., Novick D., Wallach D.;
RT "Two tumor necrosis factor-binding proteins purified from human
RT urine. Evidence for immunological cross-reactivity with cell surface
RT tumor necrosis factor receptors.";
RL J. Biol. Chem. 265:1531-1536(1990).
[6]
RX SEQUENCE OF 22-40; 65-69; 136-141; 300-306 AND 346-362.
RX MEDLINE=91056048; PubMed=2173696;
RA Loetscher H., Schlaeger E.J., Lahm H.-W., Pan Y.-C.E., Lesslauer W.,
RA Brockhaus M.;
RT "Purification and partial amino acid sequence analysis of two
RT distinct tumor necrosis factor receptors from HL60 cells.";
RL J. Biol. Chem. 265:20131-20138(1990).
[7]
RX CHARACTERIZATION.
RX MEDLINE=93016040; PubMed=1328224;
RA Pennica D., Lam V.T., Mize N.K., Weber R.F., Lewis M., Fendly B.M.,
RA Lipari M.F., Goeddel D.V.;
RT "Biochemical properties of the 75-kDa tumor necrosis factor receptor.
RT Characterization of ligand binding, internalization, and receptor
RT phosphorylation.";
RL J. Biol. Chem. 267:21172-21178(1992).
[8]
RX X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 419-428 IN TRAF2 COMPLEX.
RX MEDLINE=99221490; PubMed=10206649;
RA Park Y.C., Burkitt V., Villa A.R., Tong L., Wu H.;
RT "Structural basis for self-association and receptor recognition of
RT human TRAF2.";
RL Nature 398:533-538(1999).
CC -!- FUNCTION: RECEPTOR FOR TNF-ALPHA. HIGH AFFINITY FOR TNA-ALPHA AND
CC APPROXIMATELY 5-FOLD LOWER AFFINITY FOR TNF-BETA.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- PTM: PHOSPHORYLATED; MAINLY ON SERINE RESIDUES WITH A VERY LOW
CC LEVEL ON THREONINE RESIDUES.
CC -!- PHARMACEUTICAL: AVAILABLE UNDER THE NAME ENBREL (IMMUNEX AND
CC WYETH-AVERTS). USED TO TREAT MODERATE TO SEVERE RHEUMATOID
CC ARTHRITIS (RA). ENBREL CONSIST OF THE EXTRACELLULAR LIGAND-BINDING
CC PORTION OF TNFR2 LINKED TO AN IMMUGLOBULIN FC CHAIN. IT BINDS TO
CC TNF-ALPHA AND BLOCKS ITS INTERACTIONS WITH RECEPTORS.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD120b entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd120b.htm".
CC -!- DATABASE: NAME=Enbrel; NOTE=Clinical information on Enbrel;
CC WWW="http://www.enbrelinfo.com/".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M32315; AAC59929.1; -
DR EMBL; M35857; AAC63262.1; -
DR EMBL; U52165; AAC50622.1; -
DR EMBL; U52156; AAC50622.1; JOINED.
DR EMBL; U52157; AAC50622.1; JOINED.
DR EMBL; U52158; AAC50622.1; JOINED.
DR EMBL; U52159; AAC50622.1; JOINED.
DR EMBL; U52160; AAC50622.1; JOINED.
DR EMBL; U52161; AAC50622.1; JOINED.

DR EMBL; U52162; AAC50622.1; JOINED.
DR EMBL; U52163; AAC50622.1; JOINED.
DR EMBL; U52164; AAC50622.1; JOINED.
DR EMBL; M55994; AAA36755.1; -
DR PIR; A35356; A35356.
DR PIR; A36007; A36007.
DR PIR; A36475; A36475.
DR PIR; B35010; B35010.
DR PIR; A23666; A23666.
DR PDB; 1CA9; 12-APR-99.
DR MIM; 191191; -
DR INTERPRO; IPR001368; -
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00500; TNFR_NGFR_2; 4.
KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal;
KW Phosphorylation; Pharmaceutical; 3D-structure.
FT SIGNAL 1 22 TUMOR NECROSIS FACTOR RECEPTOR 2.
FT CHAIN 23 461 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 23 257 POTENTIAL.
FT TRANSEM 258 287 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 288 461
FT DOMAIN 39 201 4 X TNFR-CYS.
FT REPEAT 39 76 TNFR-CYS 1.
FT REPEAT 77 118 TNFR-CYS 2.
FT REPEAT 119 162 TNFR-CYS 3.
FT REPEAT 163 201 TNFR-CYS 4.
FT DISULFID 40 53 BY SIMILARITY.
FT DISULFID 54 67 BY SIMILARITY.
FT DISULFID 57 75 BY SIMILARITY.
FT DISULFID 78 93 BY SIMILARITY.
FT DISULFID 96 110 BY SIMILARITY.
FT DISULFID 100 118 BY SIMILARITY.
FT DISULFID 120 126 BY SIMILARITY.
FT DISULFID 134 143 BY SIMILARITY.
FT DISULFID 137 161 BY SIMILARITY.
FT DISULFID 164 179 BY SIMILARITY.
FT CARBOHYD 171 171 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 193 193 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 141 141 R -> P (IN REF. 4).
FT CONFLICT 196 196 R -> M (IN REF. 1 AND 3).
FT CONFLICT 363 363 A -> T (IN REF. 4).
SQ SEQUENCE 461 AA; 48316 MW; 603B580ECD67636F CRC64;

Query Match 23.0%; Score 189; DB 1; Length 461;
Best Local Similarity 33.0%; Pred. No. 1.4e-09;
Matches 34; Conservative 12; Mismatches 51; Indels 6; Gaps 1;

Qy 16 RLVCACCPGTFVQPCRRDPTTCGCPPRHYTFQWYLERCRVNLGGEREEARAC 75
Db 51 QMCSKSPGQAHKVFCTKTDVDCSDCEDSTYTQLMWNVPECLSCGSRSSDQVETQAC 110
Qy 76 HATHNRCACRTGFFAHAG-----FCLHASCPPGAGVIAPG 112
Db 111 TREQNRICTRGWYCALSKQEGCLCAPLRKCRFGVGARPG 153

RESULT 4
ID TNRC_HUMAN STANDARD; PRT; 435 AA.
AC P36941;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE LYMPHOTOXIN-BETA RECEPTOR PRECURSOR (TUMOR NECROSIS FACTOR RECEPTOR
DE 2 RELATED PROTEIN) (TUMOR NECROSIS FACTOR C RECEPTOR).
GN LTBR OR TNFR OR TNFRSF3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.

```
TISSUE=LIVER;
MEDLINE=93252381; PubMed=8486360;
Baens M., Chaffanet M., Cassiman J.J., den Berghe H., Marynen P.;
"Construction and evaluation of a hncDNA library of human l2p
transcribed sequences derived from a somatic cell hybrid.";
Genomics 16:214-218(1993).
RN
RP FUNCTION.
RX MEDLINE=94225209; PubMed=8171323;
Crowe P.D., van Arsdale T.L., Walter B.N., Ware C.F., Hession C.,
Enrenfels B., Browning J.L., Din W.S., Goodwin R.G., Smith C.A.;
"A lymphotoxin-beta-specific receptor.";
Science 264:707-710(1994).
RL
CC -!- FUNCTION: RECEPTOR FOR THE LYMPHOTOXIN-BETA. POSSIBLE FUNCTION IN
CC IMMUNE DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L04270; AAA36757.1; .
CC HSP; P25942; 1CDF.
CC MIM; 600979; .
CC INTERPRO; IPR001368; .
CC PFAM; PF00040; TNFR_C6; 4.
CC PROSITE; PS00652; TNFR_NGFR_1; 2.
CC PROSITE; PS00500; TNFR_NGFR_2; 3.
CC Receptor; Transmembrane; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 30
FT CHAIN 31 435
FT DOMAIN 31 227
FT TRANSMEM 228 248
FT DOMAIN 249 435
FT DOMAIN 42 211
FT REPEAT 42 81
FT REPEAT 82 124
FT REPEAT 125 168
FT REPEAT 169 211
FT REPEAT 43 58
FT DISULFID 59 72
FT DISULFID 62 80
FT DISULFID 83 98
FT DISULFID 101 116
FT DISULFID 104 124
FT DISULFID 126 132
FT DISULFID 139 148
FT DISULFID 142 167
FT DISULFID 170 185
FT CARBOHYD 40 40
FT CARBOHYD 177 177
FT SEQUENCE 435 AA; 46709 MW; 6246266022F656F CRC64;
SQ
Query Match 22.7%; Score 186; DB 1; Length 435;
Best Local Similarity 37.9%; Pred. No. 2.3e-09;
Matches 39; Conservative 12; Mismatches 40; Indels 12; Gaps 5;
QY 12 ETGERLVCAQCPGTFVQPCRRDSDPTTCGCPPRHYTFQFNWYLCRCRYCNVLCGEREEARACH 68
DB 52 EPOHRIICCSRCPPGTVSAKSRIRDTVCATCAENSYNEHNYLTICQLCRCPDPMVG-- 109
QY 69 EBEARACHATHNRACRRTGTFFAHAGFCLE--H-----ASCPGG 105
DB 110 LEEIAPCTSKRKTQRCQPMFC-AAWALECTHCELLSDCPGG 151
RESULT 5
```

```
VT2_SFVKA STANDARD; PRT; 325 AA.
ID VT2_SFVKA
AC P25943;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE TUMOR NECROSIS FACTOR SOLUBLE RECEPTOR PRECURSOR (PROTEIN T2).
GN T2.
OS Shope fibroma virus (strain Kasza) (SFV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Leporipoxvirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87321103; PubMed=2820128;
Upton C., Delange A.M., McFadden G.;
"Murine poxviruses: genomic organization and DNA sequence of the
telomeric region of the Shope fibroma virus genome.";
Virology 160:20-30(1987).
RL
RN [2]
RP FUNCTION.
RX MEDLINE=91207415; PubMed=1850261;
Smith C.A., Davis T., Wignall J.M., Din W.S., Farrah T., Upton C.,
McFadden G., Goodwin R.G.;
"T2 open reading frame from the Shope fibroma virus encodes a soluble
form of the TNF receptor.";
Biochem. Biophys. Res. Commun. 176:335-342(1991).
RL
CC -!- FUNCTION: BINDS TO TNF-ALPHA AND BETA. PROBABLY PREVENTS TNF TO
CC REACH CELLULAR TARGET AND THEREBY DEAMPENING THE POTENTIAL
CC ANTIVIRAL EFFECTS OF THE CYTOKINE.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M17433; .; NOT ANNOTATED_CDS.
CC EMBL; A23727; CAA01687.1; .
CC FIR; B43692; B43692.
CC HSP; F19438; ITNR.
CC INTERPRO; IPR001368; .
CC PFAM; PF00020; TNFR_C6; 2.
CC PROSITE; PS00652; TNFR_NGFR_1; 2.
CC PROSITE; PS00500; TNFR_NGFR_2; 1.
CC Receptor; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 16
FT CHAIN 17 325
FT DOMAIN 27 186
FT REPEAT 27 62
FT REPEAT 63 104
FT REPEAT 105 147
FT REPEAT 148 186
FT REPEAT 187 205
FT CARBOHYD 105 105
FT CARBOHYD 181 181
FT CARBOHYD 205 205
FT CARBOHYD 238 238
FT SEQUENCE 325 AA; 35132 MW; 810530339198A71E CRC64;
SQ
Query Match 20.8%; Score 170.5; DB 1; Length 325;
Best Local Similarity 31.4%; Pred. No. 4e-08;
Matches 37; Conservative 8; Mismatches 58; Indels 15; Gaps 2;
QY 17 LVCAQCPGTFVQPCRRDSDPTTCGCPPRHYTFQFNWYLCRCRYCNVLCGEREEARACH 76
DB 38 LCCASCHPGFYASRLCGPGSNTVCSPCDGTFTASTNHAPACVSCRGCTGHLSEQPCD 97
QY 77 ATHNRACRRTGTFFA-----HAGFCLHNASCPGAGVINAPGESWARGPSGGRR 128
DB 98 RTHDRVNCSTGNYCLKKGONGCRICAPQTKCPAGYGV-----SGHTRAGDTLC 146
```

```

RESULT 6
VT2_MXVL
AC P29825; STANDARD; PRT; 326 AA.
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE TUMOR NECROSIS FACTOR SOLUBLE RECEPTOR PRECURSOR (PROTEIN T2).
GN T2.
OS Myxoma virus (strain Lausanne).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Leporipoxvirus.
RN [1]
RX MEDLINE=91335768; PubMed=1651597;
RA Upcon C., Macen J.L., Schreiber M., McFadden G.;
RT "Myxoma virus expresses a secreted protein with homology to the tumor
RT necrosis factor receptor gene family that contributes to viral
RT virulence.";
RL Virology 184:370-382(1991).
CC -!- FUNCTION: BINDS TO TNF-ALPHA AND BETA. PROBABLY PREVENTS TNF TO
CC REACH CELLULAR TARGET AND THEREBY DEAMPENING THE POTENTIAL
CC ANTIVIRAL EFFECTS OF THE CYTOKINE.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M95181; AAA46632.1; -.
DR EMBL; A23729; CA01688.1; -.
DR PIR; A40566; GOVZML.
DR HSSP; P19438; LTNR.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 2.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00050; TNFR_NGFR_2; 2.
DR KW Receptor; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 16
FT CHAIN 17 326
FT DOMAIN 27 186
FT REPEAT 27 62
FT REPEAT 63 104
FT REPEAT 105 147
FT REPEAT 148 186
FT CARBOHYD 66 66
FT CARBOHYD 181 181
FT CARBOHYD 205 205
FT CARBOHYD 238 238
FT CARBOHYD 326 AA; 35208 MW; ABBF027E947292FF CRC64;
SQ SEQUENCE 326 AA; 35208 MW; 44956 MW; 29B326A56AEF661 CRC64;

Query Match 20.2%; Score 165.5; DB 1; Length 326;
Best Local Similarity 29.7%; Pred. No. 1.le-07;
Matches 35; Conservative 10; Mismatches 58; Indels 15; Gaps 2;

QY 17 LVCAOCPCTFVQRCRDSPTTCGCPPRHYTQFWNVLRCRYCNVLCGEREEARACH 76
DB 38 LCCTSCPPGSGYASRLCGPGSDVTCSPCKNETFTASTNHAPACVSCRGCTGHLSESQCD 97
QY 77 ATHNACRCRTGFFA-----HAGFCLEHAGCPPGAGVIAPGESNARGAPRSGRRRC 128
DB 98 KTRDRVCDGSAGNYLLKGQECRCICAPKTKCPAGYGV-----SGHRTGDVLC 146

RESULT 7
TNRC_MOUSE
AC P50284; STANDARD; PRT; 415 AA.
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE LYMPOTOXIN-BETA RECEPTOR PRECURSOR.
DE LTBOR OR TNFR.
GN GN
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RX MEDLINE=96072804; PubMed=7594541;
RA Force W.R., Walter B.N., Hession C., Tizard R., Kozak C.A.,
RA Browning J.L., Ware C.F.;
RT "Mouse lymphotoxin-beta receptor. Molecular genetics, ligand binding,
RT and expression.";
RL J. Immunol. 155:5280-5288(1995).
RN [2]
RX MEDLINE=96163885; PubMed=8586432;
RA Nakamura T., Tashiro K., Nazarea M., Nakano T., Sasayama S.,
RA Honjo T.;
RT "The murine lymphotoxin-beta receptor cDNA: isolation by the signal
RT sequence trap and chromosomal mapping.";
RL Genomics 30:312-319(1995).
CC -!- FUNCTION: RECEPTOR FOR THE LYMPOTOXIN-BETA. POSSIBLE FUNCTION IN
CC IMMUNE DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U29173; AAA68964.1; -.
DR EMBL; L38423; AAB00846.1; -.
DR EMBL; U30798; AAA81334.1; -.
DR HSSP; P25942; 1CDF.
DR MGD; MGI:104875; LTBR.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00050; TNFR_NGFR_2; 3.
DR KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 30
FT CHAIN 31 415
FT DOMAIN 31 223
FT TRANSMBM 224 244
FT DOMAIN 245 415
FT DOMAIN 42 213
FT REPEAT 42 81
FT REPEAT 82 124
FT REPEAT 125 170
FT REPEAT 171 213
FT DISULFID 43 58
FT DISULFID 59 72
FT DISULFID 62 80
FT DISULFID 83 98
FT DISULFID 101 116
FT DISULFID 104 124
FT DISULFID 126 132
FT DISULFID 139 150
FT DISULFID 142 169
FT DISULFID 172 187
FT CARBOHYD 40 40
FT CARBOHYD 179 179
FT CARBOHYD 415 AA; 44956 MW; 29B326A56AEF661 CRC64;
SQ SEQUENCE 415 AA; 44956 MW; 29B326A56AEF661 CRC64;

```

FT	DOMAIN	30	169	4 X TNFR-CYS.
FT	REPEAT	30	65	TNFR-CYS 1.
FT	REPEAT	66	107	TNFR-CYS 2.
FT	REPEAT	108	126	TNFR-CYS 3 (INCOMPLETE).
FT	REPEAT	127	167	TNFR-CYS 4.
FT	CARBOHYD	146	146	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	160	160	N-LINKED (GLCNAC. . .) (POTENTIAL).
SEQ	SEQUENCE	277 AA;	29340 MW;	49F15525941550BF CRC64;

Query Match 19.1%; Score 157; DB 1; Length 415;
 Best Local Similarity 34.0%; Pred. No. 7.le-07;
 Matches 33; Conservative 10; Mismatches 42; Indels 12; Gaps 4;

QY	19	CAQCPGTFVORPCRRDPTTCGPPRHHTYQFWNL--EERCYGNVLGCGEERARAC 75
QY	19	CAQCPGTFVORPCRRDPTTCGPPRHHTYQFWNL--EERCYGNVLGCGEERARAC 75
DB	59	CSRCPPGEFFAVACSQDTVCKTCHPNHSYNEHNLSTCQLCRPCDVLG--FEEVAPC 116
QY	76	HATHNRACRCRTGFFA-----HAGFCLEH--ASCPPG 105
DB	117	TSDRACRCQPGMSVYLDNECVHCEERLVLCQPG 153

Query Match 19.1%; Score 147; DB 1; Length 277;
 Best Local Similarity 31.9%; Pred. No. 3.7e-06;
 Matches 37; Conservative 14; Mismatches 51; Indels 14; Gaps 5;

QY	5	TYPRDAETGERLVCAQCPGTFVORPCRRDPTTCGPPRHHTYQFWNY--LERCRCYN 62
QY <th>5</th> <th>TYPRDAETGERLVCAQCPGTFVORPCRRDPTTCGPPRHHTYQFWNY--LERCRCYN 62</th>	5	TYPRDAETGERLVCAQCPGTFVORPCRRDPTTCGPPRHHTYQFWNY--LERCRCYN 62
DB <th>35</th> <th>TYPSNDR-----CCECRPGMGVSRCSRONTVCRCPGFGFYNDVSSKPKPCPTWCN 88</th>	35	TYPSNDR-----CCECRPGMGVSRCSRONTVCRCPGFGFYNDVSSKPKPCPTWCN 88
QY <th>63</th> <th>VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116</th>	63	VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116
DB <th>89</th> <th>LRSG--SERKQLCTNTQDTCRCRAGCTQPLDSYKPGVDCAPCPGHH--FSPGDNQA 140</th>	89	LRSG--SERKQLCTNTQDTCRCRAGCTQPLDSYKPGVDCAPCPGHH--FSPGDNQA 140

Query Match 17.9%; Score 147; DB 1; Length 277;
 Best Local Similarity 31.9%; Pred. No. 3.7e-06;
 Matches 37; Conservative 14; Mismatches 51; Indels 14; Gaps 5;

QY	63	VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116
QY <th>63</th> <th>VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116</th>	63	VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116
DB <th>89</th> <th>LRSG--SERKQLCTNTQDTCRCRAGCTQPLDSYKPGVDCAPCPGHH--FSPGDNQA 140</th>	89	LRSG--SERKQLCTNTQDTCRCRAGCTQPLDSYKPGVDCAPCPGHH--FSPGDNQA 140

Query Match 17.9%; Score 147; DB 1; Length 277;
 Best Local Similarity 31.9%; Pred. No. 3.7e-06;
 Matches 37; Conservative 14; Mismatches 51; Indels 14; Gaps 5;

QY	63	VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116
QY <th>63</th> <th>VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116</th>	63	VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116
DB <th>89</th> <th>LRSG--SERKQLCTNTQDTCRCRAGCTQPLDSYKPGVDCAPCPGHH--FSPGDNQA 140</th>	89	LRSG--SERKQLCTNTQDTCRCRAGCTQPLDSYKPGVDCAPCPGHH--FSPGDNQA 140

Query Match 17.9%; Score 147; DB 1; Length 277;
 Best Local Similarity 31.9%; Pred. No. 3.7e-06;
 Matches 37; Conservative 14; Mismatches 51; Indels 14; Gaps 5;

QY	63	VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116
QY <th>63</th> <th>VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116</th>	63	VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116
DB <th>89</th> <th>LRSG--SERKQLCTNTQDTCRCRAGCTQPLDSYKPGVDCAPCPGHH--FSPGDNQA 140</th>	89	LRSG--SERKQLCTNTQDTCRCRAGCTQPLDSYKPGVDCAPCPGHH--FSPGDNQA 140

Query Match 17.9%; Score 147; DB 1; Length 277;
 Best Local Similarity 31.9%; Pred. No. 3.7e-06;
 Matches 37; Conservative 14; Mismatches 51; Indels 14; Gaps 5;

QY	63	VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116
QY <th>63</th> <th>VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116</th>	63	VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116
DB <th>89</th> <th>LRSG--SERKQLCTNTQDTCRCRAGCTQPLDSYKPGVDCAPCPGHH--FSPGDNQA 140</th>	89	LRSG--SERKQLCTNTQDTCRCRAGCTQPLDSYKPGVDCAPCPGHH--FSPGDNQA 140

Query Match 17.9%; Score 147; DB 1; Length 277;
 Best Local Similarity 31.9%; Pred. No. 3.7e-06;
 Matches 37; Conservative 14; Mismatches 51; Indels 14; Gaps 5;

QY	63	VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116
QY <th>63</th> <th>VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116</th>	63	VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116
DB <th>89</th> <th>LRSG--SERKQLCTNTQDTCRCRAGCTQPLDSYKPGVDCAPCPGHH--FSPGDNQA 140</th>	89	LRSG--SERKQLCTNTQDTCRCRAGCTQPLDSYKPGVDCAPCPGHH--FSPGDNQA 140

Query Match 17.9%; Score 147; DB 1; Length 277;
 Best Local Similarity 31.9%; Pred. No. 3.7e-06;
 Matches 37; Conservative 14; Mismatches 51; Indels 14; Gaps 5;

QY	63	VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116
QY <th>63</th> <th>VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116</th>	63	VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116
DB <th>89</th> <th>LRSG--SERKQLCTNTQDTCRCRAGCTQPLDSYKPGVDCAPCPGHH--FSPGDNQA 140</th>	89	LRSG--SERKQLCTNTQDTCRCRAGCTQPLDSYKPGVDCAPCPGHH--FSPGDNQA 140

Query Match 17.9%; Score 147; DB 1; Length 277;
 Best Local Similarity 31.9%; Pred. No. 3.7e-06;
 Matches 37; Conservative 14; Mismatches 51; Indels 14; Gaps 5;

QY	63	VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116
QY <th>63</th> <th>VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116</th>	63	VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116
DB <th>89</th> <th>LRSG--SERKQLCTNTQDTCRCRAGCTQPLDSYKPGVDCAPCPGHH--FSPGDNQA 140</th>	89	LRSG--SERKQLCTNTQDTCRCRAGCTQPLDSYKPGVDCAPCPGHH--FSPGDNQA 140

Query Match 17.9%; Score 147; DB 1; Length 277;
 Best Local Similarity 31.9%; Pred. No. 3.7e-06;
 Matches 37; Conservative 14; Mismatches 51; Indels 14; Gaps 5;

QY	63	VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116
QY <th>63</th> <th>VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116</th>	63	VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116
DB <th>89</th> <th>LRSG--SERKQLCTNTQDTCRCRAGCTQPLDSYKPGVDCAPCPGHH--FSPGDNQA 140</th>	89	LRSG--SERKQLCTNTQDTCRCRAGCTQPLDSYKPGVDCAPCPGHH--FSPGDNQA 140

Query Match 17.9%; Score 147; DB 1; Length 277;
 Best Local Similarity 31.9%; Pred. No. 3.7e-06;
 Matches 37; Conservative 14; Mismatches 51; Indels 14; Gaps 5;

QY	63	VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116
QY <th>63</th> <th>VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116</th>	63	VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116
DB <th>89</th> <th>LRSG--SERKQLCTNTQDTCRCRAGCTQPLDSYKPGVDCAPCPGHH--FSPGDNQA 140</th>	89	LRSG--SERKQLCTNTQDTCRCRAGCTQPLDSYKPGVDCAPCPGHH--FSPGDNQA 140

Query Match 17.9%; Score 147; DB 1; Length 277;
 Best Local Similarity 31.9%; Pred. No. 3.7e-06;
 Matches 37; Conservative 14; Mismatches 51; Indels 14; Gaps 5;

QY	63	VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116
QY <th>63</th> <th>VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116</th>	63	VLGGEREEARACHATHNRACRCRTGFFAHAGF--CLEHASCPCGAGVIAGESWA 116
DB <th>89</th> <th>LRSG--SERKQLCTNTQDTCRCRAGCTQPLDSYKPGVDCAPCPGHH--FSPGDNQA 140</th>	89	LRSG--SERKQLCTNTQDTCRCRAGCTQPLDSYKPGVDCAPCPGHH--FSPGDNQA 140

Query Match 17.9%; Score 147; DB 1; Length 277;
 Best Local Similarity 31.9%; Pred. No. 3.7e-06

FT	CHAIN	20	271	OX40L RECEPTOR.		
FT	DOMAIN	20	210	EXTRACELLULAR (POTENTIAL).		
FT	TRANSMEM	211	235	POTENTIAL.		
FT	DOMAIN	236	271	CYTOPLASMIC (POTENTIAL).		
FT	DOMAIN	25	164	4 X TNFR-CYS.		
FT	REPEAT	25	60	TNFR-CYS 1.		
FT	REPEAT	61	102	TNFR-CYS 2.		
FT	REPEAT	103	123	TNFR-CYS 3 (INCOMPLETE).		
FT	REPEAT	124	164	TNFR-CYS 4.		
FT	CARBOHYD	143	143	N-LINKED (GLCNAC. . .) (POTENTIAL).		
SO	SEQUENCE	271 AA;	29895 MW;	C06465136B16E821 CRC64;		
Query Match 15.7%; Score 129; DB 1; Length 271;						
Best Local Similarity 30.3%; Pred. No. 0.00013;						
Matches 36; Conservative 15; Mismatches 50; Indels 18; Gaps 7;						
QY	5	TPWRDAETGERLVCAOC	PGTFFVQRCRDRSP	PTTCGCPPRHRYTQFWNY--LBRCRYCN 62		
DB	30	TPP	-----SGHK-CCRCQPGHGMVSRCDH	TRDTVCHPCPEFGFVNEAVNYDTCKQCTQC 83		
QY	63	VLGEREEERACHATN	RCRCRTGTF-----FAHAGFCL	EHASCPGAGVIAQESWA 116		
DB	84	HRSG--SELKQNTPT	EDTVCCQCRPGTQ	PRODDSHK-LGVDCVCPGCH--FSPGSNOA 137		
RESULT 11						
OX40_MOUSE STANDARD; PRT; 272 AA.						
ID	OX40_MOUSE					
AC	P47741;					
DT	01-FEB-1996	(Rel. 33, Created)				
DT	01-FEB-1996	(Rel. 33, Last sequence update)				
DE	01-OCT-2000	(Rel. 40, Last annotation update)				
DE	OX40L RECEPTOR	PRECURSOR (OX40 ANTIGEN).				
GN	TNFRSF4	OR TXGP1	OR OX40.			
OS	Mus musculus	(Mouse).				
OC	Eukaryota;	Metazoa;	Chordata;	Graniata; Vertebrata; Euteleostomi;		
OC	Mammalia;	Eutheria;	Rodentia;	Sciurognathi; Muridae; Murinae; Mus.		
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=BALB/C;					
RC	MEDLINE=94044750;	PubMed=8228223;				
RA	Calderhead D.M., Buhlmann J.E., van den Eertwegh A.J.,					
RA	Claassen E., Noelle R.J., Fell H.;					
RT	"Cloning of mouse OX40; a T cell activation marker that may mediate					
RT	T-B cell interactions";					
RL	J. Immunol. 151:5261-5271(1993).					
RN	[2]					
RP	SEQUENCE FROM N.A.					
RC	MEDLINE=95255413;	PubMed=7737295;				
RA	Birkeland M.L., Copeland N.G., Gilbert D.J., Jenkins N.A.,					
RA	Barclay A.N.;					
RT	"Gene structure and chromosomal localization of the mouse homologue					
RT	of rat OX40 protein.";					
RL	Eur. J. Immunol. 25:926-930(1995).					
CC	-1- FUNCTION: RECEPTOR FOR THE OX40L/GP34 CYTOKINE.					
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.					
CC	-1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.					
CC	-----					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation					
CC	the European Bioinformatics Institute. There are no restrictions on its					
CC	use by non-profit institutions as long as its content is in no way					
CC	modified and this statement is not removed. Usage by and for commercial					
CC	entities requires a license agreement (See http://www.isb-sib.ch/announcement					
CC	or send an email to license@isb-sib.ch).					
CC	-----					
CC	EMBL; 221674; CAA79772.1; -					
CC	EMBL; X85214; CAA59476.1; -					
DR	HSSP; P25942; ICDF.					
DR	MGI; MGI:104512; TXGP1.					
DR	INTERPRO; IPR001368; -					
DR	PIR; PFO0020; TNFR G6; 3.					


```

or send an email to license@isb-sib.ch).
-----
CC EMBL: U34794; AAC48546.1; -
CC HSSP: P25445; LDDF.
DR INTERPRO: IPR000488; -
DR INTERPRO: IPR001368; -
DR PFAM: PF00020; TNFR_G6; 3.
DR PFAM: PF00531; death; 1.
DR PROSITE: PS00652; TNFR_NGFR_1; 2.
DR PROSITE: PS00652; TNFR_NGFR_2; 2.
DR PROSITE: PS00050; TNFR_NGFR_2; 2.
DR PROSITE: PS00050; TNFR_NGFR_2; 2.
KW Apoptosis; Receptor; Glycoprotein; Transmembrane; Repeat; Signal.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 323 FASL RECEPTOR.
FT DOMAIN 17 170 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 171 188 POTENTIAL.
FT DOMAIN 189 323 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 45 163 3 X TNFR-CYS.
FT REPEAT 45 80 TNFR-CYS 1.
FT REPEAT 81 124 TNFR-CYS 2.
FT REPEAT 125 163 TNFR-CYS 3.
FT REPEAT 238 306 DEATH DOMAIN.
FT DOMAIN 323 364 4D88A90E9E1F4892 CRC64;
SQ SEQUENCE 323 AA; 36445 MW; 4D88A90E9E1F4892 CRC64;

Query Match 15.6%; Score 128; DB 1; Length 272;
Best Local Similarity 31.4%; Pred. No. 0.00016;
Matches 37; Conservative 14; Mismatches 51; Indels 16; Gaps 7;

QY 5 TYPWRDAGTGERLVAOCPPGTFVQRCRDSPTTCGCPGPRHYTOFWNY--LERCRYCN 62
DB 31 TYP-----SGHK-CCRCQPGHGWSCDHTDTLCHPCETGFYNAVNYDKQCTQCN 84
QY 63 VLCCEREERARACHATHNRCRCRTGFFAH--AGFCL--EHASCPPGAGVIAPGESWA 116
DB 85 HRSQ--SELKONCTPTDVTVCRCRPGTPQPRQDSGYKLGVDVCPGPH--FSPGNQQA 138

RESULT 12
FASA_BOVIN STANDARD; PRT; 323 AA.
AC FASA_BOVIN
AC P51867;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE FASL RECEPTOR PRECURSOR (APOPTOSIS-MEDIATING SURFACE ANTIGEN FAS)
DE (APO-1 ANTIGEN) (CD95).
GN TNFRSF6 OR APT1 OR FAS.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
RN [1]
RX MEDLINE-96226401; PubMed-86341151;
RA Yoo J., Stone R.T., Beattie C.W.;
RT "Cloning and characterization of the bovine Fas. ";
RL DNA Cell Biol. 15:227-234(1996)
CC -!- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS FASL. THE
CC ADAPTOR MOLECULE FADD RECRUITS CASPASE-8 TO THE ACTIVATED
CC RECEPTOR. THE RESULTING AGGREGATE CALLED THE DEATH-INDUCING
CC SIGNALING COMPLEX (DISC) PERFORMS CASPASE-8 PROTEOLYTIC
CC ACTIVATION. ACTIVE CASPASE-8 INITIATES THE SUBSEQUENT CASCADE OF
CC CASPASES (ASPARTATE-SPECIFIC CYSTEINE PROTEASES) MEDIATING
CC APOPTOSIS. FAS-MEDIATED APOPTOSIS MAY HAVE A ROLE IN THE
CC INDUCTION OF PERIPHERAL TOLERANCE, IN THE ANTIGEN-STIMULATED
CC SUICIDE OF MATURE T-CELLS, OR BOTH (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- DOMAIN: CONTAINS A DEATH DOMAIN INVOLVED IN THE BINDING OF FADD,
CC AND MAYBE TO OTHER CYTOSOLIC ADAPTOR PROTEINS.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -!- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation.
The European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/)

```


Wed Jan 31 13:45:09 2001

```

Query Match          15.4%; Score 126.5; DB 1; Length 260;
Best Local Similarity 32.7%; Pred. No. 0.0002;
Matches 32; Conservative 10; Mismatches 35; Indels 21; Gaps 6;

QY 16 RLVAQCPGPTGVQRPC-RRDSPTTCGCP-----PPRHYTFWNVLERCVN--VLGG 66
DB 37 KLCCMCEPCTFLVKDQDQHRKAACDPCIPGVSPFDHHTR--PHCESCRHCNSGLL-- 92
QY 67 EREERARACHATHNRCACRTGTFNAGFCLEHASCPP 104
DB 93 -----VRNCTITANAECACRNGWQCRDKECTE-----CDP 122

RESULT 15
TNRI_MOUSE STANDARD; PRT; 454 AA.
AC P25118;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR 1 PRECURSOR (P60) (TNF-R1) (TNF-R1)
DE (P55).
GN TNFRSF1A OR TNFR1 OR TNFR-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-91187885; PubMed-1849278;
RA Lewis M., Tagaglia L.A., Lee A., Bennett G.L., Rice G.C.,
RA Wong G.H., Chen E.Y., Goeddel D.V.;
RT "Cloning and expression of cDNAs for two distinct murine tumor
RT necrosis factor receptors demonstrate one receptor is species
RT specific."
RL Proc. Natl. Acad. Sci. U.S.A. 88:2830-2834(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-91246168; PubMed-1645445;
RA Goodwin R.G., Anderson D., Jerzy R., Davis T., Brannan C.I.,
RA Copeland N.G., Jenkins N.A., Smith C.A.;
RT "Molecular cloning and expression of the type 1 and type 2 murine
RT receptors for tumor necrosis factor."
RL Mol. Cell. Biol. 11:3020-3026(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE-91285014; PubMed-1647956;
RA Barrett K., Taylor-Fishwick D.A., Cope A.P., Kissoneghis A.M.,
RA Gray P.W., Feldmann M., Foxwell B.M.J.;
RT "Cloning, expression and cross-linking analysis of the murine p55
RT tumor necrosis factor receptor."
RL Eur. J. Immunol. 21:1649-1656(1991).
RN [4]
RP SEQUENCE FROM N.A.
RX TISSUE-SPLEEN;
RX MEDLINE-92039815; PubMed-1657766;
RA Rothe J.G., Brockhaus M., Gentz R., Lesslauer W.;
RT "Molecular cloning and expression of the mouse Tnf receptor type b."
RL Immunogenetics 34:338-340(1991).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE-94245292; PubMed-8188324;
RA Bebo B.F., Linthicum D.S.;
RT "Nucleotide sequence of the TNF type I receptor from a mouse
RT endothelioma cell line."
RL Immunogenetics 39:450-451(1994).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE-93156721; PubMed-8381516;
RA Rothe J., Bluetmann H., Gentz R., Lesslauer W., Steinmetz M.;
RT "Genomic organization and promoter function of the murine tumor
RT necrosis factor receptor beta gene."
RL Mol. Immunol. 30:165-175(1993).

```

```

-|- FUNCTION: RECEPTOR FOR TNF-ALPHA. THE ADAPTOR MOLECULE FADD
CC RECRUITS CASPASE-8 TO THE ACTIVATED RECEPTOR. THE RESULTING
CC AGGREGATE CALLED THE DEATH-INDUCING SIGNALING COMPLEX (DISC)
CC PERFORMS CASPASE-8 PROTEOLYTIC ACTIVATION WHICH INITIATES THE
CC SUBSEQUENT CASCADE OF CASPASES (ASPARTATE-SPECIFIC CYSTEINE
CC PROTEASES) MEDIATING APOPTOSIS (BY SIMILARITY).
CC -|- SUBUNIT: TNF BINDING TO THE EXTRACELLULAR DOMAIN OF TNFR1 LEADS TO
CC HOMOPRIMERIZATION. ONCE AGGREGATED THE RECEPTORS SPECIFICALLY
CC PROVIDE A NOVEL MOLECULAR INTERFACE THAT INTERACTS SPECIFICALLY
CC WITH THE DEATH DOMAIN OF TRADD. VARIOUS TRADD-INTERACTING
CC PROTEINS SUCH AS TRAFs, RIP AND POSSIBLY FADD, ARE RECRUITED TO
CC TNFR1 COMPLEX BY THEIR ASSOCIATION WITH TRADD. THIS COMPLEX
CC ACTIVATES AT LEAST TWO DISTINCT SIGNALING CASCADES, APOPTOSIS AND
CC NF-KAPPA B SIGNALING (BY SIMILARITY).
CC -|- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -|- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -|- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M60468; AAA39751.1; -
CC EMBL: M59377; AAA40464.1; -
CC EMBL: X59238; CAM1922.1; -
CC EMBL: X57796; CAA40936.1; -
CC EMBL: L26349; AAA59361.1; -
CC EMBL: M76656; AAA40465.1; -
CC EMBL: M8067; AAA40465.1; JOINED.
CC EMBL: M76655; AAA40465.1; JOINED.
CC PIR: A38634; GOMST1.
CC PIR: S16677; S16677.
CC PIR: S19021; S19021.
CC HSSP: P19438; TEXT.
CC MGD: MGI:1314884; TNFRSF1A.
CC INTERPRO: IPR000488; -
CC INTERPRO: IPR001368; -
CC PFAM: PF00020; TNFR_C6; 4.
CC PFAM: PF00531; death; 1
CC PROSITE: PS00652; TNFR_NGFR_1; 3.
CC PROSITE: PS50050; TNFR_NGFR_2; 3.
CC PROSITE: PS50017; DEATH_DOMAIN; 1.
CC Receptor; Transmembrane; Glycoprotein; Repeat; Signal; Apoptosis.
KW SIGNAL 1 21
FT CHAIN 22 454
FT DOMAIN 22 212
FT TRANSMEM 213 235
FT DOMAIN 236 454
FT DOMAIN 43 196
FT REPEAT 43 82
FT REPEAT 83 125
FT REPEAT 126 166
FT REPEAT 167 196
FT DOMAIN 339 349
FT DOMAIN 356 441
FT DISULFID 44 58
FT DISULFID 59 72
FT DISULFID 62 81
FT DISULFID 84 99
FT DISULFID 102 117
FT DISULFID 105 125
FT DISULFID 127 143
FT DISULFID 146 158
FT DISULFID 149 166
FT DISULFID 168 179
FT DISULFID 182 191
FT DISULFID 185 195
FT CARBOHYD 54 54
FT CARBOHYD 151 151

```

FT CARBOHYD 202 202 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 394 394 R -> G (IN REF. 6).
SQ SEQUENCE 454 AA; 50129 MW; 0710C2E8C3C2B6D9 CRC64;

Query Match 15.3%; Score 125.5; DB 1; Length 454;
Best Local Similarity 22.6%; Pred. No. 0.00039;
Matches 35; Conservative 14; Mismatches 47; Indels 59; Gaps 5;

Qy 17 LVCAQCPGTFVQPCRDSP---TTCGCPPRHYTQFWNYLRCRYC----- 61
Db 57 ICCTKCHKGTLYLVSDC--PSPGRDVCRECEKGTFTASQNYLRQCLCKCKEMSQVEI 114
Qy 62 -----NVLGGEREEEAR-----ACHATHNRACRGTGFFA 91
Db 115 SPCQADKDTVCCKENQFQRYLSETHFCVDCSPCFNGTVTIPCKETONTVCNCHAGEFL 174
Qy 92 HAGFCLHASC-----PPGAGVIAPGES 114
Db 175 RESECVPCSHCKKNECKMKLCLPPPLANVTNPQDS 209

Search completed: January 30, 2001, 17:06:08
Job time: 1229 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:54:56 ; Search time 254.27 Seconds
(without alignments)
64.534 Million cell updates/sec

Title: US-09-518-931-4_COPY_31_170
Perfect score: 821
Sequence: 1 AETPTVPWRDAETGERLVCA.....PRSGRRRCRGQVAGPSLAP 140

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_15:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	667	81.2	300	4	095407
2	292.5	35.6	372	4	Q9UHP4
3	292.5	35.6	401	4	O00300
4	285.5	34.8	401	11	O08712
5	282.5	34.4	401	11	O08727
6	230.5	28.1	302	13	Q9PU50
7	206	25.1	459	11	O62327
8	199.5	24.3	482	11	O08734
9	196.5	24.2	349	12	O57100
10	196.5	23.9	348	12	O57277
11	196.5	23.9	348	12	O57103
12	196.5	23.9	348	12	O57108
13	196.5	23.9	349	12	O57291
14	196.5	23.9	349	12	O57101
15	196.5	23.9	349	12	O57102
16	195.5	23.8	349	12	O57099
17	191.5	23.3	348	12	O57112
18	191.5	23.3	348	12	O85407
19	191.5	23.3	349	12	O57110

20	191.5	23.3	349	12	O57111
21	191.5	23.3	349	12	O89118
22	191.5	23.3	349	12	O89098
23	190	23.1	349	12	O57284
24	190	23.1	349	12	O57098
25	189	23.0	439	4	Q16042
26	188	22.9	326	12	O57120
27	187.5	22.8	655	4	O75509
28	186	22.7	349	12	O57097
29	185	22.5	326	12	O57122
30	185	22.5	351	12	O57117
31	185	22.5	351	12	O73559
32	184	22.4	350	12	O57123
33	184	22.4	355	12	O85308
34	183.5	22.4	349	12	O57109
35	183.5	22.4	350	12	O57116
36	182.5	22.2	347	12	O57119
37	182.5	22.2	360	12	O57118
38	179	21.8	349	12	O57305
39	178	21.7	351	12	O57121
40	175	21.3	347	12	O57115
41	166	20.2	322	12	O72761
42	164	20.0	320	12	O57079
43	160	19.5	316	12	O57092
44	160	19.5	320	12	O57091
45	160	19.5	320	12	O57300

ALIGNMENTS

RESULT 1

ID	O95407	PRELIMINARY:	PRT;	300 AA.
AC	O95407;			
DT	01-MAY-1999	(TREMBLrel. 10, Created)		
DT	01-MAY-1999	(TREMBLrel. 10, Last sequence update)		
DT	01-OCT-2000	(TREMBLrel. 15, Last annotation update)		
DE	DECOY RECEPTOR 3 (M68)	(M68C) (M68E)		
GN	DCR3 OR TR6.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID:9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=95087326; PubMed=9872321;			
RA	Pitt R.M., Marsters S.A., Lawrence D.A., Roy M., Kischkel F.C.,			
RA	Dowd P., Huang A., Donahue C.J., Sherwood S.W., Baldwin D.T.,			
RA	Godowski P.J., Wood W.I., Gurney A.L., Hillan K.J., Cohen R.L.,			
RA	Godard A.D., Botstein D., Ashkenazi A.			
RT	"Genomic amplification of a decoy receptor for Fas ligand in lung and			
RT	colon cancer";			
RL	Nature 396:699-703(1998).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=BLOOD;			
RE	MEDLINE=99253915; PubMed=10318773;			
RA	Yu K.Y., Kwon B., Ni J., Zhai Y., Ebner R., Kwon B.S.;			
RT	"A newly identified member of tumor necrosis factor receptor			
RT	superfamily (TR6) suppresses LIGHT-mediated apoptosis.";			
RL	J. Biol. Chem. 274:13733-13736(1999).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=PANCREAS;			
RX	MEDLINE=2012600; PubMed=10655513;			
RA	Bai C., Connolly B., Metzger M.L., Hilliard C.A., Liu X., Sandig V.,			
RA	Soderman A., Galloway S.M., Liu Q., Austin C.P., Caskey C.T.;			
RT	"Overexpression of M68/Dcr3 in human gastrointestinal tract tumors			
RT	independent of gene amplification and its location in a four-gene			
RT	cluster.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 97:1230-1235(2000).			
DR	EMBL; AF104419; AAD03056.1; .			

Wed Jan 31 13:45:10 2001

```

DR EMBL: AF134240; AAD29688.1; -
DR EMBL: AF217796; AAF35244.1; -
DR EMBL: AF217793; AAF33685.1; -
DR EMBL: AF217794; AAF33686.1; -
DR HSSP: P25942; 1CDF.
DR INTERPRO: IPR000561; -
DR INTERPRO: IPR001388; -
DR PFAM: PF000020; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE: PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
DR PROSITE: PS50050; TNFR_NGFR_2; 2.
DR PRODOM: PD000771; -; 1.
KW Receptor. 300 AA; 32679 MW; F90AEE33718449AF CRC64;
SQ SEQUENCE 300 AA; 32679 MW; F90AEE33718449AF CRC64;

Query Match 81.2%; Score 667; DB 4; Length 300;
Best Local Similarity 100.0%; Pred. No. 1.7e-59;
Matches 112; Conservative 0; Mismatches 0; Gaps 0;

Qy 1 AETTPWRDAETGERLVCAQCPTGVQVPCRRDSTTCGCPPPRHHTQFWNYLRCRY 60
Dd 31 AETTPWRDAETGERLVCAQCPTGVQVPCRRDSTTCGCPPPRHHTQFWNYLRCRY 90
Qy 61 CNVLGGEREEARACHATHNRACRRTGFFAHAGFCLEHASCPGAGVIAPG 112
Dd 91 CNVLGGEREEARACHATHNRACRRTGFFAHAGFCLEHASCPGAGVIAPG 142

RESULT 2
Q9UHP4 PRELIMINARY; PRT; 372 AA.
AC Q9UHP4;
DT 01-MAY-2000 (TRENBLrel. 13; Created)
DT 01-OCT-2000 (TRENBLrel. 13; Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15; Last annotation update)
DE OSTEOPROTEGERIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=98351569; Pubmed=9688283;
RA Morinaga T., Nakagawa N., Yasuda H., Tsuda E., Higashio K.;
RA "Cloning and characterization of the gene encoding human
RA osteoprotegerin/osteoclastogenesis-inhibitory factor.";
RL Eur. J. Biochem. 254:685-691(1998).
DR EMBL: AF134187; AAF20168.1; -
DR HSSP: P25942; 1CDF.
DR INTERPRO: IPR001368; -
DR PFAM: PF000020; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE: PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE: PS50050; TNFR_NGFR_2; 2.
DR NON_TER 1
FT SEQUENCE 372 AA; 42758 MW; F02527A5CD01CCD3 CRC64;

Query Match 35.6%; Score 292.5; DB 4; Length 372;
Best Local Similarity 38.5%; Pred. No. 7.7e-22;
Matches 50; Conservative 21; Mismatches 50; Indels 9; Gaps 2;

Qy 4 PTPWRDAETGERLVCAQCPTGVQVPCRRDSTTCGCPPPRHHTQFWNYLRCRY 63
Dd 5 PKYLYDEETSHQLCDKCPGPIYGLQHCFTAKWKVCAPCPDHYTDSWHTSDECLYCS 64
Qy 64 LCGEREERACHATHNRACRRTGFFAHAGFCLEHASCPGAGVIAPGSGWARGAPRS 123
Dd 65 VKELQYVQKBCNTHNRVCEKRGYLETEFCLKHKRSCPPGGV-----QAGTPER 117
Qy 124 GG--RCGRG 131
Dd 118 NTVCKRCPDG 127

```

```

RESULT 3
Q00300 PRELIMINARY; PRT; 401 AA.
AC Q00300; O60236;
DT 01-JUL-1997 (TRENBLrel. 04; Created)
DT 01-JUL-1997 (TRENBLrel. 04; Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15; Last annotation update)
DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)
DE (OCIF) (TUMOR NECROSIS FACTOR RECEPTOR SUPERFAMILY MEMBER 11B).
GN TNFRSF11B OR OPG OR OCIF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RX MEDLINE=97262071; Pubmed=9108485;
RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,
RA Luethy R., Nguyen H.O., Wooden S., Bennett L., Boone T., Shimamoto G.,
RA Derose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,
RA Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,
RA Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,
RA Sugds S., Boyle W.J.;
RA "Osteoprotegerin: a novel secreted protein involved in the regulation
RA of bone density.";
RL Cell 89:309-319(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG FIBROBLAST;
RX MEDLINE=98151033; Pubmed=9492069;
RA Yasuda H., Shima N., Nakagawa N., Mochizuki S.-I., Yano K., Fujise N.,
RA Sato Y., Goto M., Yamaguchi K., Kuriyama M., Kanno T., Murakami A.,
RA Tsuda E., Morinaga T., Higashio K.;
RA "Identity of osteoclastogenesis inhibitory factor (OCIF) and
RA osteoprotegerin (OPG): a mechanism by which OPG/OCIF inhibits
RA osteoclastogenesis in vitro.";
RL Endocrinology 139:1329-1337(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=98351569; Pubmed=9688283;
RA Morinaga T., Nakagawa N., Yasuda H., Tsuda E., Higashio K.;
RA "Cloning and characterization of the gene encoding human
RA osteoprotegerin/osteoclastogenesis-inhibitory factor.";
RL Eur. J. Biochem. 254:685-691(1998).
CC -1- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES
CC OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY
CC SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO
CC OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN
CC STROMAL CELLS AND OSTEOCLAST PROGENITORS.
CC -1- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM) (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGHEST LEVELS IN LUNG, HEART,
CC KIDNEY, PLACENTA, THYROID, SPINAL CORD AND LIVER. ALSO DETECTED IN
CC A NUMBER OF OTHER HEMATOPOIETIC AND IMMUNE ORGANS. NOT DETECTED IN
CC THE PANCREAS OR PERIPHERAL BLOOD LYMPHOCYTES.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC EMBL: AB002146; BAA25910.1; -
CC EMBL: AB008822; BAA32076.1; -
CC EMBL: AB008821; BAA32076.1; JOINED.
CC EMBL: U94332; AAB53709.1; -
CC HSSP: P25942; 1CDF.
CC MIN: 602643; -
CC INTERPRO: IPR001368; -
CC DR PFAM: PF00020; TNFR_NGFR_1; UNKNOWN_1.
CC DR PROSITE: PS00652; TNFR_NGFR_1; UNKNOWN_1.
CC DR PROSITE: PS50050; TNFR_NGFR_2; 2.
CC DR PRODOM: PD000771; -; 1.
CC DR Glycoprotein; Repeat; Cytokine; Signal.
CC KW SIGNAL 1 21 BY SIMILARITY.
FT

```

```

FT CHAIN 22 401 OSTEOPROTEGERIN.
FT DOMAIN 23 183 4 X TNFR-CYS.
FT REPEAT 23 63 TNFR-CYS 1.
FT REPEAT 64 106 TNFR-CYS 2.
FT REPEAT 107 143 TNFR-CYS 3.
FT REPEAT 144 201 TNFR-CYS 4.
FT DOMAIN 306 365 DEATH DOMAIN.
FT DISULFID 41 54 BY SIMILARITY.
FT DISULFID 44 62 BY SIMILARITY.
FT DISULFID 65 80 BY SIMILARITY.
FT DISULFID 83 97 BY SIMILARITY.
FT DISULFID 87 105 BY SIMILARITY.
FT DISULFID 118 142 BY SIMILARITY.
FT DISULFID 145 160 BY SIMILARITY.
FT CARBOHYD 98 98 POTENTIAL.
FT CARBOHYD 152 152 POTENTIAL.
FT CARBOHYD 165 165 POTENTIAL.
FT CARBOHYD 178 178 POTENTIAL.
FT CARBOHYD 289 289 POTENTIAL.
FT CONFLICT 263 263 A -> D (IN REF. 2 AND 3).
SQ SEQUENCE 401 AA; 45996 MW; EB42FA51C9D7C71E CRC64;

Query Match 35.6%; Score 292.5; DB 4; Length 401;
Best Local Similarity 38.5%; Pred. No. 8.3e-22;
Matches 50; Conservative 21; Mismatches 50; Indels 9; Gaps 2;

Qy 4 PTYPWRDAETGERLYCAOCPGTFVORPCRRDSPPTTCGCPPRHYTQWNYLERCRYCNV 63
Db 26 PKYLHYDETSQHLCDKCPCTGTYLKQHTAKWTKVCAPCPDHYVYDTSWHTSDCYLCS 85

Qy 64 LCGEREERACHATHNACRCRTGFFAHAGFLEHASCPCPGAGVIAPGESWARGAPRS 123
Db 86 VKELQYVKQCNTRHNRCVCEKRGYLEIEFCLKHSRCPGFGV-----QAGTPEP 138

Qy 124 GG--RRCGRG 131
Db 139 NTVCKRCPDG 148

RESULT 4
O08712 PRELIMINARY; PRT; 401 AA.
ID O08712; O70202;
AC 01-JUL-1997 (Tremblrel. 04, Created)
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)
DE (OCIF).
GN TNFRSF11B OR OPG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=KIDNEY;
RA Simonet W.S.; Lacey D.L.; Dunstan C.R.; Kelley M.; Chang M.-S.;
RA Luethy R.; Nguyen H.Q.; Woodson S.; Bennett L.; Boone T.; Shimamoto G.;
RA Derose M.; Elliott R.; Colombero A.; Tan H.-L.; Trail G.; Sullivan J.;
RA Davy E.; Bucay N.; Renshaw-Gegg L.; Hughes T.M.; Hill D.; Pattison W.;
RA Campbell P.; Sander S.; Van G.; Tarpley J.; Derby P.; Lee R.;
RA Sugis S.; Boyle W.J.;
RT "Osteoprotegerin: a novel secreted protein involved in the regulation
of bone density."
RL Cell 89:309-319(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/OLA, AND NIH SWISS;
RX MEDLINE=98382527; PubMed=9714833;
RA Mizuno A.; Murakami A.; Nakagawa N.; Yasuda H.; Tsuda E.; Morinaga T.;
RA Higashio K.;

```

```

RT "Structure of the mouse osteoclastogenesis inhibitory factor (OCIF)
gene and its expression in embryogenesis.";
RL Gene 215:339-343(1998).
CC -!- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES
OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY
SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO
OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN
STROMAL CELLS AND OSTEOCLAST PROGENITORS.
CC -!- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM).
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -!- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN LIVER, LUNG,
BRAIN, HEART, KIDNEY, STOMACH, INTESTINE, SKIN, CALVARIA AND
PLACENTA. NOT DETECTED IN SPLEEN.
CC -!- DEVELOPMENTAL STAGE: IN THE EMBRYO, HIGH LEVELS ARE DETECTED AT
DAY 7. AT DAY 11, EXPRESSION DECREASES AND THEN INCREASES FROM DAY
15 TO DAY 17.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
EMBL; U94331; AAB53708.1; -.
DR EMBL; AB013898; BAA28269.1; -.
DR EMBL; AB013903; BAA33388.1; -.
DR EMBL; AB013899; BAA33388.1; JOINED.
DR EMBL; AB013900; BAA33388.1; JOINED.
DR EMBL; AB013901; BAA33388.1; JOINED.
DR EMBL; AB013902; BAA33388.1; JOINED.
DR HSSP; P25942; 1CDF.
DR MGD; MGI:109587; Opg.
DR INTERPRO; IPR000488; -.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS00117; DEATH_DOMAIN; 1.
DR PROSITE; PS00500; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -. 1
KW Glycoprotein; Repeat; Cytokine; Signal.
FT SIGNAL 1 21
FT CHAIN 22 401 OSTEOPROTEGERIN.
FT DOMAIN 23 201 4 X TNFR-CYS.
FT REPEAT 23 63 TNFR-CYS 1.
FT REPEAT 64 106 TNFR-CYS 2.
FT REPEAT 107 143 TNFR-CYS 3.
FT REPEAT 144 201 TNFR-CYS 4.
FT DOMAIN 306 365 DEATH DOMAIN.
FT DISULFID 41 54 BY SIMILARITY.
FT DISULFID 44 62 BY SIMILARITY.
FT DISULFID 65 80 BY SIMILARITY.
FT DISULFID 83 97 BY SIMILARITY.
FT DISULFID 87 105 BY SIMILARITY.
FT DISULFID 118 142 BY SIMILARITY.
FT DISULFID 145 160 BY SIMILARITY.
FT CARBOHYD 98 98 POTENTIAL.
FT CARBOHYD 165 165 POTENTIAL.
FT CARBOHYD 178 178 POTENTIAL.
FT CARBOHYD 289 289 POTENTIAL.
FT VARIANT 138 138 R -> P (IN STRAINS 129/OLA AND NIH
SWISS).
FT VARIANT 161 161 I -> R (IN STRAINS 129/OLA AND NIH
SWISS).
FT VARIANT 165 165 N -> D (IN STRAINS 129/OLA AND NIH
SWISS).
FT VARIANT 288 288 S -> A (IN STRAINS 129/OLA AND NIH
SWISS).
FT VARIANT 296 296 L -> R (IN STRAINS 129/OLA AND NIH
SWISS).
SQ SEQUENCE 401 AA; 45923 MW; CAA6102D3B312470 CRC64;

Query Match 34.8%; Score 285.5; DB 11; Length 401;
Best Local Similarity 37.6%; Pred. No. 4.2e-21;
Matches 53; Conservative 22; Mismatches 55; Indels 11; Gaps 3;

Qy 4 PTYPWRDAETGERLYCAOCPGTFVORPCRRDSPPTTCGCPPRHYTQWNYLERCRYCNV 63
Db 26 PKYLHYDETSQHLCDKCPCTGTYLKQHTAKWTKVCAPCPDHYVYDTSWHTSDCYLCS 85

```

```

SQ SEQUENCE 401 AA; 46192 MW; FEC6A31F1D4E573A CRC64;

Query Match 34.4%; Score 282.5; DB 11; Length 401;
Best Local Similarity 38.3%; Pred. No. 8.3e-21; Indels 11; Gaps 3;
Matches 54; Conservative 20; Mismatches 56;

QY 4 PTPWRDAETGERLVCAOCPGTFVQRCRRDSTTCGPPRRHYTFQWNYLERCRYCNV 63
Db 26 PKYLHYDPETGRQLLCKDKCAPGYLKHQCTVRRKTLVCPGPDYSYDTSWHTSDECVYCS 85
QY 64 LCGEREERARACHATHNRACRCRTGFFAHAGFCLEHACGPPGAGYIAPGESWARGAPRS 123
Db 86 VKELQTVKQECNTRNRVCECEGRYLEFCLKHRSCPPGLGLV-----QAGTPER 138

QY 124 GG--RRCGRGOVAG--PSLAP 140
Db 139 NTVCKKCPDGFPSGETSSKAP 159

RESULT 6
Q9PUS0 PRELIMINARY; PRT; 302 AA.
AC Q9PUS0:
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE DECOY RECEPTOR.
OS Salvelinus fontinalis (Brook trout).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
ON NCBI_TaxID=8038;
RX [1]
RP SEQUENCE FROM N.A.
RA Bobe J., Goetz F.W.;
RT "A tumor necrosis factor receptor homolog is up-regulated in the brook
trout (Salvelinus fontinalis) ovary at the completion of ovulation.";
RL Biol. Reprod. 0:0-0(1999).
DR EMBL; AF156738; AAD56428.1; -.
DR HSSP; PI9438; 1EXT.
DR INTERPRO; IPR000561; -.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS01866; EGF_2; UNKNOWN_1.
DR PROSITE; PS50050; TNFR_NGFR_2; 1.
KW Receptor.
SQ SEQUENCE 302 AA; 34037 MW; E44C73477F05C3DF CRC64;

Query Match 28.1%; Score 230.5; DB 13; Length 302;
Best Local Similarity 33.6%; Pred. No. 1.1e-15;
Matches 37; Conservative 25; Mismatches 45; Indels 3; Gaps 2;

QY 5 TYPWRDAETGERLVCAOCPGTFVQRCRRDSTTCGPPRRHYTFQWNYLERCRYCNV 64
Db 22 TFKNDTRYGLSLVCDRCPPGTYLRAPCSAMRKSQCAECPCNGAYTEFNHHSKLRCS-M 80
QY 65 CGEREERARACHATHNRACRCRTGFFAHAGF--CLEHACGPPGAGYIAPG 112
Db 81 CAENQVVKQECSPSNCECECKEGYFNKKYEACIKHKCECPGPGYANTTG 130

RESULT 7
Q62327 PRELIMINARY; PRT; 459 AA.
AC Q62327:
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR 2 MRNA (FRAGMENT).
OS Mus musculus (Mouse).
```


OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN NCBI_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=NOD;
RX MEDLINE=95178848; PubMed=7873884;
RA Powell E.E., Wicker L.S., Peterson L.B., Todd J.A.;
RT "Allelic variation of the type 2 tumor necrosis factor receptor
gene";
RL Mamm. Genome 5:726-727(1994).
DR EMBL; X76401; CAA53981.1; -;
DR HSSP; P19438; INCF.
DR INTERPRO; IPR001368; -;
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 3.
FT NON_TER 1
FT VARIANT 87 87 S -> T.
FT VARIANT 93 93 T -> I.
FT VARIANT 268 268 F -> I.
FT VARIANT 345 345 S -> F.
FT VARIANT 421 421 Y -> C.
SQ SEQUENCE 459 AA; 48686 MW; 6C51D2CF1C4626DF CRC64;

Query Match 25.18; Score 206; DB 11; Length 459;
Best Local Similarity 31.88; Pred. No. 4.4e-13; Indels 22; Gaps 3;
Matches 42; Conservative 13; Mismatches 55;

Qy 16 RLVCAQCPGGTFVQPCRRDPTTCGPPRRHYTFQWNYLRCRYCNVLCGEREEARAC 75
: : ||||| : : : : ||||| : : : : ||||| : : : : |||||
Db 37 QMCCAKCPGGQVVRHFCNKTSDTVCADCEASMYTQVWVNFRTCLSCSSCSTDQVETRA 96
: : ||||| : : : : ||||| : : : : ||||| : : : : |||||

Qy 76 HATHNRACRRTGFF----AHAGF---CLEHASCPGAGVIAPGESWARGGRRRC 128
: : ||||| : : : : ||||| : : : : ||||| : : : : |||||
Db 97 TKQONRVACAGRYCALKTHSGSCRCQMLSKCGPGFV-----ASSRA 141
: : ||||| : : : : ||||| : : : : ||||| : : : : |||||

Qy 129 GRGOVAGPSLAP 140
: : || : : ||
Db 142 PNGNVLCACAP 153
: : || : : ||

RESULT 8
O88734 PRELIMINARY; PRT; 482 AA.
ID O88734;
AC O88734;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE P80 TNF-ALPHA RECEPTOR.
GN TNFR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN NCBI_TaxID=10090;
RP SEQUENCE FROM N.A.
RA Hurie B., Segade F., Rodriguez R., Ramos S.S., Lazo P.S.;
RT "The Mouse Tumor Necrosis Factor Receptor Gene: Genomic Structure and
Characterization of the two Transcripts.";
RL Genomics 0:0-0(0).
DR EMBL; Y14619; CAA74969.1; -;
DR EMBL; Y14620; CAA74969.1; JOINED.
DR EMBL; Y14621; CAA74969.1; JOINED.
DR EMBL; Y14622; CAA74969.1; JOINED.
DR EMBL; Y14623; CAA74969.1; JOINED.
DR EMBL; Y14624; CAA74969.1; JOINED.
DR HSSP; P19438; INCF.
DR INTERPRO; IPR001368; -;
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 3.

DR PRODOM; PD000771; -; 1.
SQ SEQUENCE 482 AA; 51106 MW; F6C15046B48FF83C CRC64;

Query Match 24.38; Score 199.5; DB 11; Length 482;
Best Local Similarity 30.98; Pred. No. 2.1e-12;
Matches 43; Conservative 13; Mismatches 54; Indels 29; Gaps 4;

Qy 16 RLVCAQCPGGTFVQPCRR-----DSPTTGCPGPPRRHYTFQWNYLRCRYCNVLCGER 68
: : ||||| : : : : ||||| : : : : ||||| : : : : |||||
Db 52 QMCCAKCPGGQVVRHFCNKTSDTVCADCEASMYTQVWVNFRTCLSCSSCSTD 111
: : ||||| : : : : ||||| : : : : ||||| : : : : |||||
Qy 69 EEEARACHATHNRACRRTGFF----AHAGF---CLEHASCPGAGVIAPGESWARGGAP 121
: : ||||| : : : : ||||| : : : : ||||| : : : : |||||
Db 112 QVETRACTKQONRVACAGRYCALKTHSGSCRCQMLSKCGPGFV----- 158
: : ||||| : : : : ||||| : : : : ||||| : : : : |||||
Qy 122 RSGGRCRGGOVAGPSLAP 140
: : || : : ||
Db 159 --ASSRAPNGNVLCACAP 175
: : || : : ||

RESULT 9
O57100 PRELIMINARY; PRT; 349 AA.
ID O57100;
AC O57100;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN CRMB.
OS Monkeypox virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10244;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIGERIA-1971 (71-0082);
RA Loparev V.N., Parsons J.M., Esposito J.J.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U87844; AAB94361.1; -;
DR HSSP; P25942; ICDF.
DR INTERPRO; IPR001368; -;
DR PFAM; PF00020; TNFR_C6; 2.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -; 1.
SQ SEQUENCE 349 AA; 38239 MW; DF6C280D478F2422 CRC64;

Query Match 24.28; Score 198.5; DB 12; Length 349;
Best Local Similarity 35.58; Pred. No. 2e-12;
Matches 38; Conservative 13; Mismatches 49; Indels 7; Gaps 2;

Qy 9 RDAETGER-LVCAQCPGGTFVQPCRRDPTTCGPPRRHYTFQWNYLRCRYCNVLCGE 67
: : ||||| : : : : ||||| : : : : ||||| : : : : |||||
Db 33 KDNEYRSNLCLSCPPGTYASRLCDSKTQCTPCGSDTFTSHNHLQACLSGCRGDS 92
: : ||||| : : : : ||||| : : : : ||||| : : : : |||||
Qy 68 REEEARACHATHNRACRRTGFF-----AHAGFCLHASCPGAGV 108
: : ||||| : : : : ||||| : : : : ||||| : : : : |||||
Db 93 NQVETRSCNTTHNRICEGSPGYCLLKAGSGCSTCKGIGYGV 139
: : ||||| : : : : ||||| : : : : ||||| : : : : |||||

RESULT 10
O57277 PRELIMINARY; PRT; 348 AA.
ID O57277;
AC O57277;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN CRMB.
OS Monkeypox virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;

[illegible]

[illegible]

```
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 2.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00650; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -; 1.
SQ SEQUENCE 349 AA; 36308 MW; CBD2C949F994C59C CRC64;

Query Match      23.9%; Score 196.5; DB 12; Length 349;
Best Local Similarity 35.5%; Pred. No. 3.1e-12;
Matches 38; Conservative 13; Mismatches 49; Indels 7; Gaps 2;

QY   9 RDAETGER-LVCAOCPGGTFVQRPCRDSPTTCGPCPRHYTQFWNVYLERCRYCNVLGE 67
    :| |: | ||||| |: | ||||| |: | ||||| |: | |||||
Db   33 KDNEYRSNLNCLSCPPGYASRLKDSKTITQTCTPCGSDFTTSHNNHQLACLSCNGRCD 92
    :| |: | ||||| |: | ||||| |: | ||||| |: | |||||

QY   68 REEARACHATHNRACRCRTGFF-----AHAGFCLEHASCPPGAGV 108
    :| |: | ||||| |: | ||||| |: | ||||| |: | |||||
Db   93 NOVETRCSNTTHNRICECSPGYYLLKGASGCRFTCIISKTKGIGYGV 139
    :| |: | ||||| |: | ||||| |: | ||||| |: | |||||

Search completed: January 30, 2001, 16:54:57
Job time: 913 sec
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:47:46 ; Search time 132.77 seconds
(without alignments)
18.935 Million cell updates/sec

Title: US-09-518-931-4_COPY_31_170
Perfect score: 821
Sequence: 1 AETPTVPWRDAETGERLVCA.....PRSGRRRCRGQVAGPSLAP 140

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues
Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/PTUS.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	667	81.2	300	2	US-08-794-796-2	Sequence 2, Appli
2	292.5	35.6	401	3	US-08-974-022-6	Sequence 6, Appli
3	285.5	34.8	401	3	US-08-974-022-4	Sequence 4, Appli
4	282.5	34.4	401	3	US-08-974-022-2	Sequence 2, Appli
5	203	24.7	474	2	US-08-650-000-4	Sequence 4, Appli
6	203	24.7	474	5	5395760-4	Patent No. 5395760
7	191	23.3	163	2	US-08-219-237B-5	Sequence 5, Appli
8	189	23.0	164	2	US-08-232-087A-9	Sequence 9, Appli
9	189	23.0	227	3	US-08-974-022-48	Sequence 48, Appli
10	189	23.0	461	1	US-08-385-229-2	Sequence 2, Appli
11	189	23.0	461	2	US-08-650-000-2	Sequence 2, Appli
12	189	23.0	461	5	5395760-2	Patent No. 5395760
13	189	23.0	486	1	US-08-243-010-1	Sequence 1, Appli
14	189	23.0	518	1	US-08-385-229-4	Sequence 4, Appli
15	187.5	22.8	655	3	US-08-959-382-2	Sequence 2, Appli
16	186	22.7	197	2	US-08-505-606-1	Sequence 1, Appli
17	184	22.4	355	1	US-08-292-549-6	Sequence 6, Appli
18	170.5	20.8	207	3	US-08-974-022-47	Sequence 47, Appli
19	170.5	20.8	325	1	US-08-292-549-2	Sequence 2, Appli
20	170.5	20.8	325	4	PCT-US91-02207-2	Sequence 2, Appli
21	165.5	20.2	326	1	US-08-292-549-4	Sequence 4, Appli
22	165.5	20.2	326	4	PCT-US91-02207-4	Sequence 4, Appli
23	163	19.9	120	3	US-08-974-022-42	Sequence 42, Appli
24	147	17.9	277	2	US-08-147-784-2	Sequence 2, Appli
25	143.5	17.5	283	4	PCT-US96-12374-2	Sequence 2, Appli
26	140.5	17.1	186	1	US-08-089-458B-6	Sequence 6, Appli
27	134	16.3	122	2	US-08-232-087A-7	Sequence 7, Appli
28	129.5	15.8	139	2	US-08-219-237B-8	Sequence 8, Appli

Sequence 4, Appli
Sequence 2, Appli
Sequence 6, Appli
Sequence 51, Appli
Sequence 7, Appli
Sequence 11, Appli
Sequence 11, Appli
Sequence 15, Appli
Sequence 4, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 41, Appli

ALIGNMENTS

RESULT 1
US-08-794-796-2
; Sequence 2, Application US/08794796
; Patent No. 5885800
; GENERAL INFORMATION:
; APPLICANT: Emery, John
; APPLICANT: Tan, KB
; APPLICANT: Truneh, Alem
; APPLICANT: Young, Peter
; TITLE OF INVENTION: Tumor Necrosis Related Receptor,
; TITLE OF INVENTION: TRA
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/794.796
; FILING DATE: 04-FEB-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: GH50000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-4026
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 300 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-794-796-2

Query Match 81.2%; Score 667; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 5.1e-54;

Db 52 QMCCAKCPGQVYKFCNKTSDVACDEASMYTQVWQFRTCLSCSSCTTDQVETAC 111
Qy 76 HATHNRACRCRTGFF-----AHAGF---CLEHASCPPGAGVIAPGESWARGAPRSGGRR 128
Db 112 TKOQNRVCAEAGRYCALKTHSGSCRQCMRLSKCGPGFGV-----ASSRA 156
Qy 129 GRQVAGPSLAP 140
Db 157 PNGNVCKACAP 168

RESULT 7
US-08-219-237B-5
; Sequence 5, Application US/08219237B
; Patent No. 5874546
; GENERAL INFORMATION:
; APPLICANT: NAGATA, Shigekazu
; APPLICANT: ITOH, Naoto
; APPLICANT: YONEHARA, Shin
; TITLE OF INVENTION: DNA Coding for Human Cell Surface Antigen
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James W. Hellwege
; STREET: P.O. Box 2266 Eads Station
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/219,237B
; FILING DATE: 28-MAR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/872,129
; FILING DATE: 22-APR-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: James W. Hellwege
; REGISTRATION NUMBER: 28,808
; REFERENCE/DOCKET NUMBER: 516762
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 163 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-219-237B-5

Query Match 23.3%; Score 191; DB 2; Length 163;
Best Local Similarity 33.0%; Pred. No. 9.8e-11;
Matches 34; Conservative 13; Mismatches 50; Indels 6; Gaps 1;
Qy 16 RLVCACQPPGTFVQRCRDSPPTCGPPRHVTQFWNYLERCRCYNVLCGEREEARAC 75
Db 13 QMCCSKCSPGOHAKVFCTKTSOTVCDSCEDSYTQLWNVPECLSCGSRCSDDQVETQAC 72
Qy 76 HATHNRACRCRTGFFAHAG-----FCLHASCPPGAGVIAPG 112
Db 73 TREQNRCTCRPGWYCALSKQEGCRCLCAPLRKCRPGFGVARPG 115

RESULT 8
US-08-232-087A-9
; Sequence 9, Application US/08232087A
; Patent No. 5866372
; GENERAL INFORMATION:
; APPLICANT: Stein, Harald

APPLICANT: D r kop, Horst
APPLICANT: Latza, Ute
TITLE OF INVENTION: Lymphoid CD30-Antigen
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22042

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,087A
FILING DATE: 08-SEP-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 756-103P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 164 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
FEATURE:
NAME/KEY: Protein
LOCATION: 1..164
OTHER INFORMATION: /note= "TNFR2, see Fig. 5"
US-08-232-087A-9

Query Match 23.0%; Score 189; DB 2; Length 164;
Best Local Similarity 33.0%; Pred. No. 1.5e-10;
Matches 34; Conservative 12; Mismatches 51; Indels 6; Gaps 1;
Qy 16 RLVCACQPPGTFVQRCRDSPPTCGPPRHVTQFWNYLERCRCYNVLCGEREEARAC 75
Db 14 QMCCSKCSPGOHAKVFCTKTSOTVCDSCEDSYTQLWNVPECLSCGSRCSDDQVETQAC 73
Qy 76 HATHNRACRCRTGFFAHAG-----FCLHASCPPGAGVIAPG 112
Db 74 TREQNRCTCRPGWYCALSKQEGCRCLCAPLRKCRPGFGVARPG 116

RESULT 9
US-08-974-022-48
; Sequence 48, Application US/08974022
; Patent No. 6015936
; GENERAL INFORMATION:
; APPLICANT: Boyle, William J.
; APPLICANT: Lacey, David L.
; APPLICANT: Calzone, Frank J.
; APPLICANT: Chang, Ming-Shi
; TITLE OF INVENTION: OSTEOPROTEGERIN
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California

COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,022
FILING DATE: 12-DEC-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/577,788
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-378
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 227 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-974-022-48

Query Match 23.0%; Score 189; DB 3; Length 227;
Best Local Similarity 33.0%; Pred. No. 2.1e-10;
Matches 34; Conservative 12; Mismatches 51; Indels 6; Gaps 1;
Qy 16 RLVCAQCPGFTFVORPCRRDSTTCGPPRRHYTFQFNNYLCRYCNVLCGEREEARAC 75
Db 51 QMCGSKCSPGQGHAKVFCFKTSDTVCDSCEDSTYTQLMNNVPECLSCGSRCSDDQVETQAC 110
Qy 76 HATHNRACRGTGFFAHAG-----FCLEHASCPGAGVIAPG 112
Db 111 TREQNRICTCRPGWYCALSKQEGCRLCAPLRKCRPGFGVARPG 153

RESULT 10
US-08-385-229-2
Sequence 2, Application US/08385229
Patent No. 5605690
GENERAL INFORMATION:
APPLICANT: Jacobs, Cindy A.
APPLICANT: Smith, Craig A.
TITLE OF INVENTION: Method of Treating TNF-Dependent
TITLE OF INVENTION: Inflammation Using Tumor Necrosis Factor Antagonists
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: Washington
COUNTRY: U.S.A.
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/385,229
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/946,236
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Wight, Christopher L.
REGISTRATION NUMBER: 31,680
REFERENCE/DOCKET NUMBER: 2503

TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 587-0606
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 461 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-385-229-2
Query Match 23.0%; Score 189; DB 1; Length 461;
Best Local Similarity 33.0%; Pred. No. 4.3e-10;
Matches 34; Conservative 12; Mismatches 51; Indels 6; Gaps 1;
Qy 16 RLVCAQCPGFTFVORPCRRDSTTCGPPRRHYTFQFNNYLCRYCNVLCGEREEARAC 75
Db 51 QMCGSKCSPGQGHAKVFCFKTSDTVCDSCEDSTYTQLMNNVPECLSCGSRCSDDQVETQAC 110
Qy 76 HATHNRACRGTGFFAHAG-----FCLEHASCPGAGVIAPG 112
Db 111 TREQNRICTCRPGWYCALSKQEGCRLCAPLRKCRPGFGVARPG 153
RESULT 11
US-08-650-000-2
Sequence 2, Application US/08650000
Patent No. 5945397
GENERAL INFORMATION:
APPLICANT: Smith, Craig A.
APPLICANT: Goodwin, Raymond G.
APPLICANT: Beckmann, M. Patricia
TITLE OF INVENTION: Tumor Necrosis Factor Receptors
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: Washington
COUNTRY: U.S.A.
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/650,000
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/468,453
FILING DATE:
APPLICATION NUMBER: US/08/038,765
FILING DATE:
APPLICATION NUMBER: US 403,241
FILING DATE: 05-SEP-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 405,370
FILING DATE: 11-SEP-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 421,417
FILING DATE: 13-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 523,635
FILING DATE: 10-MAY-1990
ATTORNEY/AGENT INFORMATION:
NAME: Wight, Christopher L.
REGISTRATION NUMBER: 31,680
REFERENCE/DOCKET NUMBER: 2501-D
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430

us-09-518-931-4_copy_31_170.rai

Wed Jan 31 13:45:07 2001

```
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 461 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-650-000-2

Query Match      23.0%; Score 189; DB 2; Length 461;
Best Local Similarity 33.0%; Pred. No. 4.3e-10;
Matches 34; Conservative 12; Mismatches 51; Indels 6; Gaps 1;

QY 16 RLVCAQCPPTGTVQPCRRDSTTCGPPRHYTFQWNLERCRCYCNVLCGEREEARAC 75
DB 51 QMCCSKCSPGQHAQVFKTKTSDTVCDSCDSTYTQLMNVVPECLSCGSRCSDDQVETQAC 110
QY 76 HATHNRACRRTGFFAHAG-----FCLHASCPPPGAGVIAPG 112
DB 111 TREQNRICTRPGWYCALSKQEGRCRLCAPLRCRPGFGVARPG 153

RESULT 12
5395760-2
; PATENT NO. 5395760
; APPLICANT: SMITH, CRAIG A.; GOODWIN, RAYMOND G.; BECKMANN,
; M. PATRICIA
; TITLE OF INVENTION: DNA ENCODING TUMOR NECROSIS FACTOR-a AND
; B-RECEPTORS
; NUMBER OF SEQUENCES: 17
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/523,635
; FILING DATE: 10-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 421,417
; FILING DATE: 13-OCT-1989
; APPLICATION NUMBER: 405,370
; FILING DATE: 11-SEP-1989
; APPLICATION NUMBER: 403,241
; FILING DATE: 05-SEP-1989
; SEQ ID NO: 2:
; LENGTH: 461
; 5395760-2

Query Match      23.0%; Score 189; DB 5; Length 461;
Best Local Similarity 33.0%; Pred. No. 4.3e-10;
Matches 34; Conservative 12; Mismatches 51; Indels 6; Gaps 1;

QY 16 RLVCAQCPPTGTVQPCRRDSTTCGPPRHYTFQWNLERCRCYCNVLCGEREEARAC 75
DB 51 QMCCSKCSPGQHAQVFKTKTSDTVCDSCDSTYTQLMNVVPECLSCGSRCSDDQVETQAC 110
QY 76 HATHNRACRRTGFFAHAG-----FCLHASCPPPGAGVIAPG 112
DB 111 TREQNRICTRPGWYCALSKQEGRCRLCAPLRCRPGFGVARPG 153

RESULT 13
US-08-243-010-1
; Sequence 1, Application US/08243010
; Patent No. 5639597
; GENERAL INFORMATION:
; APPLICANT: Lauffer, Leander
; APPLICANT: Zettlmeissel, Gerd
; APPLICANT: Oquendo, Patricia
; TITLE OF INVENTION: Cell-free Receptor Binding Assays, The
; TITLE OF INVENTION: Production and Use Thereof
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
```

```
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/243,010
FILING DATE: 13-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/798,564
FILING DATE: 26-NOV-1991
APPLICATION NUMBER: DE P 40 37 837.3
FILING DATE: 28-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Einaudi, Carol P.
REGISTRATION NUMBER: 32,220
REFERENCE/DOCKET NUMBER: 02481-1132-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 486 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-243-010-1

Query Match      23.0%; Score 189; DB 1; Length 486;
Best Local Similarity 33.0%; Pred. No. 4.5e-10;
Matches 34; Conservative 12; Mismatches 51; Indels 6; Gaps 1;

QY 16 RLVCAQCPPTGTVQPCRRDSTTCGPPRHYTFQWNLERCRCYCNVLCGEREEARAC 75
DB 51 QMCCSKCSPGQHAQVFKTKTSDTVCDSCDSTYTQLMNVVPECLSCGSRCSDDQVETQAC 110
QY 76 HATHNRACRRTGFFAHAG-----FCLHASCPPPGAGVIAPG 112
DB 111 TREQNRICTRPGWYCALSKQEGRCRLCAPLRCRPGFGVARPG 153

RESULT 14
US-08-385-229-4
; Sequence 4, Application US/08385229
; Patent No. 5605690
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Cindy A.
; APPLICANT: Smith, Craig A.
; TITLE OF INVENTION: Method of Treating TNF-Dependent
; TITLE OF INVENTION: Inflammation Using Tumor Necrosis Factor Antagonists
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/385,229
FILING DATE:
```

CLASSIFICATION: 435
PRIOR APPLICATION DATA: US/07/946,236
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Wight, Christopher L.
REGISTRATION NUMBER: 31,680
REFERENCE/DOCKET NUMBER: 2503
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 587-0606
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 518 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-385-229-4

Query Match 23.0%; Score 189; DB 1; Length 518;
Best Local Similarity 33.0%; Pred. No. 4.8e-10;
Matches 34; Conservative 12; Mismatches 51; Indels 6; Gaps 1;
Qy 16 RLVCAQCPPGTFVQRPCCRDSPTTCGPPRRHYTQFWNYLRCRYCNVLCGEREEARAC 75
Db 80 QMCCSKSPGQHAQKVCFTKTSDTVDCSDSTYTQLWNWPECLSCGSRCSDDQVETQAC 139
Qy 76 HATHNRACRRTGFFAHAG-----FCLEHASCPPGAGVIAPG 112
Db 140 TREQNRICRCPGWYCALSKQEGCRLCAPLRKRCRPGFGVARPG 182

RESULT 15
US-08-959-382-2
Sequence 2, Application US/08959382
Patent No. 6013476
GENERAL INFORMATION:
APPLICANT: DEEN, KEITH
APPLICANT: HURLE, MARK
APPLICANT: YOUNG, PETER
APPLICANT: TAN, K.B.
TITLE OF INVENTION: Tumor Necrosis Related Receptor,
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: RAYNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/959,382.
FILING DATE: 28-OCT-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/041,796
FILING DATE: 02-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-50017
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 655 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-959-382-2
Query Match 22.8%; Score 187.5; DB 3; Length 655;
Best Local Similarity 31.9%; Pred. No. 8.4e-10;
Matches 43; Conservative 11; Mismatches 60; Indels 21; Gaps 2;
Qy 5 TYPWRDAETGERLYVCAQCPCPGTGFVQRPCCRDSPTTCGPPRRHYTQFWNYLRCRYCNVL 64
Db 53 TYRHVDRAATGQVLTCDCPCAGTYVSEHCTNTSLRVCSGPCVGTFTTRHENGIEKCHDCSQP 112
Qy 65 CGEREEEARACHATHNRACRRTGFFAHAGFCLEHASCPPGAGVIAPGESWARGGAPRSG 124
Db 113 CPWPMIEKLPCAALTDRECTCPGPMFQSNATCAPHTVCPVGWGV-----RKK 159
Qy 125 G-----RCGRG 131
Db 160 GTETEDVRCKQCARG 174

Search completed: January 30, 2001, 16:47:47
Job time: 623 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:35:40 : Search time 183.67 seconds
(without alignments)
55.851 Million cell updates/sec

Title: US-09-518-931-2

Perfect score: 1634

Sequence: 1 MRALGPGSLCLLVLALPA.....RVARMPLERSVRERELPVH 300

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_36:*

- 1: /SIDS1/gcgdata/geneseq/geneseq/AA1980.DAT:*
- 2: /SIDS1/gcgdata/geneseq/geneseq/AA1981.DAT:*
- 3: /SIDS1/gcgdata/geneseq/geneseq/AA1982.DAT:*
- 4: /SIDS1/gcgdata/geneseq/geneseq/AA1983.DAT:*
- 5: /SIDS1/gcgdata/geneseq/geneseq/AA1984.DAT:*
- 6: /SIDS1/gcgdata/geneseq/geneseq/AA1985.DAT:*
- 7: /SIDS1/gcgdata/geneseq/geneseq/AA1986.DAT:*
- 8: /SIDS1/gcgdata/geneseq/geneseq/AA1987.DAT:*
- 9: /SIDS1/gcgdata/geneseq/geneseq/AA1988.DAT:*
- 10: /SIDS1/gcgdata/geneseq/geneseq/AA1989.DAT:*
- 11: /SIDS1/gcgdata/geneseq/geneseq/AA1990.DAT:*
- 12: /SIDS1/gcgdata/geneseq/geneseq/AA1991.DAT:*
- 13: /SIDS1/gcgdata/geneseq/geneseq/AA1992.DAT:*
- 14: /SIDS1/gcgdata/geneseq/geneseq/AA1993.DAT:*
- 15: /SIDS1/gcgdata/geneseq/geneseq/AA1994.DAT:*
- 16: /SIDS1/gcgdata/geneseq/geneseq/AA1995.DAT:*
- 17: /SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT:*
- 18: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT:*
- 19: /SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT:*
- 20: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT:*
- 21: /SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1634	100.0	300	19 W66102	Amino acid sequenc
2	1634	100.0	300	19 W63622	Human tumour necro
3	1634	100.0	300	20 Y03099	Human lung TNF-rec
4	1634	100.0	300	20 Y42182	Human FLINT #1 pro
5	1634	100.0	300	20 Y17479	Mammalian tumour n
6	1634	100.0	300	20 Y06817	Human DCR3 polypep
7	1634	100.0	300	20 W97749	Human tumour necro
8	1634	100.0	300	20 W95082	Orphan receptor (H
9	1620	99.1	300	21 Y77458	Human TNF receptor
10	1610	98.5	302	20 Y42183	Human FLINT #2 pro
11	1491	91.2	271	20 Y42184	Human mFLINT #1 pr
12	1467	89.8	273	20 Y42185	Human mFLINT #2 pr

13	1362	83.4	245	20 Y28449	A human tumour nec
14	1153	70.6	215	20 W93585	Human hAPO6 protei
15	841	51.5	153	20 Y22222	Human TNFR superfa
16	814	49.8	170	19 W63623	Human tumour necro
17	465.5	28.5	326	17 R99940	Mutated OCIF, OCIF
18	447.5	27.4	349	20 W83928	Human FTHMA-070 pa
19	447	27.4	401	17 R99931	Mutated OCIF, OCIF
20	446	27.3	401	17 R99933	Mutated OCIF, OCIF
21	445	27.2	272	17 R99944	Mutated OCIF, OCIF
22	445	27.2	321	17 R99949	Mutated OCIF, OCIF
23	445	27.2	327	17 R99941	Mutated OCIF, OCIF
24	445	27.2	351	17 R99943	Mutated OCIF, OCIF
25	445	27.2	393	17 R99948	Mutated OCIF, OCIF
26	445	27.2	399	17 R99942	Mutated OCIF, OCIF
27	445	27.2	401	17 R99934	Mutated OCIF, OCIF
28	445	27.2	401	17 R99925	Full length osteoc
29	445	27.2	401	19 W53239	Human OCIF genome
30	444	27.2	390	17 R99357	Human tumour necro
31	444	27.2	395	19 W57636	Modified TR1 recep
32	444	27.2	401	18 W38345	Human osteoprotege
33	444	27.2	401	20 Y05742	Tumour necrosis fa
34	444	27.2	401	20 W95030	Tumour necrosis fa
35	444	27.2	401	20 W83926	Human FTHMA-070 pr
36	444	27.2	401	21 Y43400	Osteoprotegerin pro
37	440.5	27.0	380	17 R99924	Mature osteoclasto
38	440.5	27.0	391	19 W53238	Human OCIF genome
39	440	26.9	401	19 W57635	TR1 receptor prote
40	438	26.8	401	17 R99932	Mutated OCIF, OCIF
41	437.5	26.8	187	17 R99950	Mutated OCIF, OCIF
42	437.5	26.8	197	17 R99945	Mutated OCIF, OCIF
43	436.5	26.7	187	21 Y77464	Primate protein se
44	435	26.6	401	17 R99935	Mutated OCIF, OCIF
45	425.5	26.0	401	18 W38343	Rat osteoprotegeri

ALIGNMENTS

RESULT 1

W66102
ID W66102 standard; Protein; 300 AA.

XX W66102;

XX 02-DEC-1998 (first entry)

XX Amino acid sequence of tumour necrosis related receptor (TR4).

XX Human; tumour necrosis related receptor; TR4; agonist; antagonist;

KW inhibition; chronic; acute; inflammation; arthritis; septicemia;

KW autoimmune disease; transplant rejection; stroke; cancer;

KW Alzheimer's disease.

XX Homo sapiens.

XX EP861850-A1.

XX 02-SEP-1998.

XX 20-JAN-1998; 98EP-0300382.

XX 04-FEB-1997; 97US-0794796.

XX (SMIK) SMITHKLINE BEECHAM CORP.

XX Emery J, Tan KB, Truneh A, Young PR;

XX WPI: 1998-508248/44.

XX N-PSDB: V07654.

XX New DNA encoding tumour necrosis related receptor - used to treat
PT and prevent e.g. inflammation, arthritis, septicemia, autoimmune
PT diseases, transplant rejection, infection, stroke, ischaemia, ARDS,

PT restenosis, AIDS, bone disorders and cancer
XX Claim 1; Fig 1; 21pp; English.
PS This is the amino acid sequence of the human tumour necrosis related
XX receptor (TR4), used in the method of the invention. The TR4 protein
CC or its agonist can be used to treat a subject in need of enhanced
CC TR4 polypeptide activity. The antagonist is used to inhibit TR4
CC polypeptide activity. The active agents can be used for the
CC treatment and prevention of diseases such as chronic and acute
CC inflammation, arthritis, septicemia, autoimmune diseases, transplant
CC rejection, stroke, cancer, Alzheimer's disease.
XX
SQ Sequence 300 AA;

Query Match 100.0%; Score 1634; DB 19; Length 300;
Best Local Similarity 100.0%; Pred. No. 3.8e-122;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALEGSLCLLVLPVAVRGVAETPTYPWRDAETGERLVCAQCPPTGVQR 60
DB 1 mralegpplsllcllvpavrgvaetpypwrdaetgerlvcaqcpptgvqr 60

QY 61 PCRRDSTTCGCPPRHYTOFWNYLRCRYCNVLCGEREEARACHATHNACRCRTGFF 120
DB 61 pcrrdspttcgpcprhytqfwnylercrycnvlgereearachathnacrctgff 120

QY 121 AHAGFCLHASCPPGAGVIAPGTPSONTCQPCPPTFSASSSSQCPHNRCTALGLA 180
DB 121 ahagfclhascpvgagviapgtpsontcqpcpptfsassssqcpnhnrctalgla 180

QY 181 LNVPGSSSDTLCTSGTGFPLSTRVFGAECERAVIDFAVDISIKRLQLQALEAPE 240
DB 181 lnvpghsshdltctsgtgfplstrvfgaeceravidfavidisikrlqlqaleape 240

QY 241 GWGPTPRAGRAALQLKRLRRITELLGAQDGLLVRLQLQALRVAMPGLERSVRERFLPVH 300
DB 241 gwgptpragraalqlkrrlritellgaqdgallvrlqlqalrvampglersvrerflpvh 300

RESULT 2
ID W63622 standard; Protein; 300 AA.
XX W63622;
AC W63622;
DT 26-OCT-1998 (first entry)
DE Human tumour necrosis factor receptor-6 alpha protein.
KW Human tumour necrosis factor receptor-6 alpha; TNFR-6 alpha; TNFR-6 beta;
KW endothelial cells; keratinocytes; normal prostate; apoptosis;
KW prostate tumour tissue.
XX Homo sapiens.
XX Key Location/Qualifiers
FT Peptide 1..30
FT Protein 31..300
FT /note= "TNFR-6 alpha"
FT Region 31..282
FT /note= "Soluble extracellular domain"
XX W09830694-A2.
XX 16-JUL-1998.
XX 13-JAN-1998; 98WO-US00153.
XX 14-JAN-1997; 97US-0035496.
XX (HUMA-) HUMAN GENOME SCI INC.
PA

XX Ebner R, Feng P, Gentz RL, Ni J, Ruben SM, Yu G;
PI WPI; 1998-399142/34.
XX N-PSDB; V39085.
XX Human tumour necrosis factor receptors 6-alpha and 6-beta - used in
PT the diagnosis of immune system-related disorder(s)
XX Claim 20; Fig 1; 91pp; English.
XX The present sequence represents the human tumour necrosis factor
CC receptor-6 alpha (TNFR-6 alpha) protein. The invention also provides
CC for the TNFR-6 beta protein (W63623). TNFR-6 alpha and TNFR-6 beta are
CC members of the tumour necrosis factor receptor (TNFR) family. TNFRs
CC are expressed in endothelial cells, keratinocytes, normal prostate and
CC prostate tumour tissue. For a number of disorders of these cells,
CC particularly of the immune system, substantially altered (whether
CC increased or decreased) levels of TNFR-6 alpha and/or TNFR-6 beta gene
CC expression can be detected, therefore the TNFR-6 alpha and TNFR-6 beta
CC polypeptides, nucleic acids and antibodies are claimed to be useful in
CC the diagnosis of such disorders. Mutations of the TNFR-6 alpha and
CC TNFR-6 beta genes can also be detected. The TNFR polypeptides are
CC also claimed to be useful for identifying ligands which may be useful
CC in the treatment of apoptosis related disorders.
XX
SQ Sequence 300 AA;

Query Match 100.0%; Score 1634; DB 19; Length 300;
Best Local Similarity 100.0%; Pred. No. 3.8e-122;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALEGSLCLLVLPVAVRGVAETPTYPWRDAETGERLVCAQCPPTGVQR 60
DB 1 mralegpplsllcllvpavrgvaetpypwrdaetgerlvcaqcpptgvqr 60

QY 61 PCRRDSTTCGCPPRHYTOFWNYLRCRYCNVLCGEREEARACHATHNACRCRTGFF 120
DB 61 pcrrdspttcgpcprhytqfwnylercrycnvlgereearachathnacrctgff 120

QY 121 AHAGFCLHASCPPGAGVIAPGTPSONTCQPCPPTFSASSSSQCPHNRCTALGLA 180
DB 121 ahagfclhascpvgagviapgtpsontcqpcpptfsassssqcpnhnrctalgla 180

QY 181 LNVPGSSSDTLCTSGTGFPLSTRVFGAECERAVIDFAVDISIKRLQLQALEAPE 240
DB 181 lnvpghsshdltctsgtgfplstrvfgaeceravidfavidisikrlqlqaleape 240

QY 241 GWGPTPRAGRAALQLKRLRRITELLGAQDGLLVRLQLQALRVAMPGLERSVRERFLPVH 300
DB 241 gwgptpragraalqlkrrlritellgaqdgallvrlqlqalrvampglersvrerflpvh 300

RESULT 3
ID Y03099 standard; Protein; 300 AA.
XX Y03099;
AC Y03099;
DT 09-DEC-1999 (first entry)
DE Human lung TNF-receptor protein.
XX Tumour necrosis factor; TNF; TNF receptor; human; lung; gene therapy;
KW detection; immunoassay; diagnosis; disease; immune system; tumour;
KW osteogenic system; cardiovascular system; central nervous system; asthma;
KW peripheral nervous systems; transplant incompatibility; antitumor;
KW rheumatoid arthritis; antiasthmatic; antiarthritic.
XX Homo sapiens.
XX OS
XX Key Location/Qualifiers
FH

QY 1 MRALEGPGLSLLCLVLALPALLPVPVAVRGVAETPTYPWRDAETGERLVCACQCPGTFVQR 60
Db 1 mralepggisllclvlalpallpvpavrgvaetptypwrdaetgerlvcacqcpptfvqr 60
QY 61 PCRRDSPTCGPPCPRHRYTQFWNYLERCYCNVLCGEREEERACHATHNRACRCRTGFF 120
Db 61 perrdspttcgpcpprhrytqfwnylercycnvlcgerereearachathnrcrcrtgff 120
QY 121 AHAGFCLEHASCPGAGVIAPTGTPSNTQCCPCCPPTFSASSSSSQCPHNRCTALGLA 180
Db 121 ahagfclehascpaggviaptgtpsgntqccppgtfsassssseqcphnrctalgla 180
QY 181 LNVPGSSSHDTLCTSGTGPPLSTRVPGAECERAVIDFVAFQDISIKRLQRLQALEAPE 240
Db 181 lnpvgssshdtlctsgtgpplstrvpgaeceravidfvafqdisikrlqrlqaleape 240
QY 241 GWGPTPRAGRAALQKLRRRLTELLGAQDQALLVRLQLALRVARMPLGERSVRERFLPVH 300
Db 241 gwgptpragraalqklrrrltellgaqdqallvrlqlalrvarmplgervrerflpvh 300

RESULT 5

Y17479
ID Y17479 standard; Protein; 300 AA.
XX Y17479;
AC
XX
DT 02-AUG-1999 (first entry)
XX
DE Mammalian tumour necrosis factor receptor OPG-2.
XX
KW Tumour necrosis factor receptor; TNF receptor; OPG-2; Paget's disease;
KW osteopenic disorder; osteoclast activity; primary osteoporosis;
KW hyperglycaemia; osteolytic metastasis; immune response; cancer.
XX
OS Mammalia.
PN WO9926977-A1.
XX
PD 03-JUN-1999.
XX
PF 24-NOV-1998; 98WO-US25065.
XX
PR 17-FEB-1998; 98US-0074896.
PR 24-NOV-1997; 97US-0066446.
XX
PA (BIOJ) BIOGEN INC.
XX
XX Tschoep J;
PI
XX
XX
XX WPI; 1999-347693/29.
XX N-PSDB; X76052.
XX
XX New tumour necrosis factor family receptor OPG-2
XX
XX Claim 1; Page 18; 22pp; English.

The present sequence represents a mammalian tumour necrosis factor receptor, designated OPG-2. OPG-2, is a member of the tumour necrosis factor receptor family, and can be used: (i) to raise specific antibodies (Ab), (ii) to treat osteopenic disorders associated with excessive osteoclast activity, e.g. primary osteoporosis, Paget's disease, hyperglycaemia of malignancy, or osteolytic metastases; (iii) for affinity purification of cognate ligands, and (iv) to screen for ligands (antagonists or agonists), Ab, or other OPG-2 blocking agents such as soluble forms of the protein, are used to prevent, or reduce severity of, an immune response, and for treating cancer. They can also be used in diagnostic assays. The nucleic acid sequence encoding OPG-2 can be used as a probe to isolate related sequences from other species.

Sequence 300 AA;

Query Match 100.0%; Score 1634; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 3.8e-122;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRALEGPGLSLLCLVLALPALLPVPVAVRGVAETPTYPWRDAETGERLVCACQCPGTFVQR 60
Db 1 mralepggisllclvlalpallpvpavrgvaetptypwrdaetgerlvcacqcpptfvqr 60
QY 61 PCRRDSPTCGPPCPRHRYTQFWNYLERCYCNVLCGEREEERACHATHNRACRCRTGFF 120
Db 61 perrdspttcgpcpprhrytqfwnylercycnvlcgerereearachathnrcrcrtgff 120
QY 121 AHAGFCLEHASCPGAGVIAPTGTPSNTQCCPCCPPTFSASSSSSQCPHNRCTALGLA 180
Db 121 ahagfclehascpaggviaptgtpsgntqccppgtfsassssseqcphnrctalgla 180
QY 181 LNVPGSSSHDTLCTSGTGPPLSTRVPGAECERAVIDFVAFQDISIKRLQRLQALEAPE 240
Db 181 lnpvgssshdtlctsgtgpplstrvpgaeceravidfvafqdisikrlqrlqaleape 240
QY 241 GWGPTPRAGRAALQKLRRRLTELLGAQDQALLVRLQLALRVARMPLGERSVRERFLPVH 300
Db 241 gwgptpragraalqklrrrltellgaqdqallvrlqlalrvarmplgervrerflpvh 300
RESULT 6
Y06817
ID Y06817 standard; Protein; 300 AA.
XX Y06817;
XX
DT 24-JUN-1999 (first entry)
XX
DE Human DcR3 polypeptide.
XX
KW DcR3 polypeptide; tumour necrosis factor receptor; TNFR; Fas ligand;
KW apoptosis; T cell mediated immune response; allergy; asthma; cancer;
KW rheumatoid arthritis; Crohn's disease; guest vs. host disease; human;
XX
XX gene therapy.
XX
XX Homo sapiens.
XX
XX WO9914330-A1.
XX
XX 25-MAR-1999.
XX
XX 18-SEP-1998; 98WO-US19661.
XX
XX 30-JUL-1998; 98US-0094640.
XX 18-SEP-1997; 97US-0059288.
XX
XX (GETH) GENENTECH INC.
XX
XX Ashkenazi AJ, Botstein D, Dodge KH, Goddard A, Gurney AL;
PI Kim KJ, Lawrence DA, Pitti R, Roy MA, Tumas DB;
PI Wood WI;
XX
XX WPI; 1999-244032/20.
XX N-PSDB; X32744.
XX
XX DcR3 polypeptide related to tumor necrosis factor receptor
XX
XX Claim 5; Fig 1; 89pp; English.
XX
XX This represents a human DcR3 polypeptide, a homologue of tumour necrosis factor receptor (TNFR) polypeptide. Host cells containing a vector comprising the DcR3 nucleic acid can be used for the recombinant expression of the protein. DcR3 binds to Fas ligand, so it (or its chimera) are useful for modulating apoptosis in mammalian cells, also other Fas-ligand induced activities, particularly to inhibit T cell mediated immune responses, e.g. in treatment of allergy, asthma, rheumatoid arthritis, Crohn's disease, guest vs. host disease etc. DcR3 may also be used to identify specific binding proteins, potential

CC inhibitors. Antibodies against DcR3 are used to treat cancer,
CC specifically of the lung and colon, also in diagnosis and for affinity
CC purification of the protein. Detecting mutations in the gene for DcR3 is
CC also used to diagnose cancer, or predisposition to it. DcR3 nucleic acid
CC is useful as hybridization probe to detect genomic or related sequences;
CC for chromosome and gene mapping; as source of antisense sequences; for
CC expression of recombinant DcR3 and to generate transgenic animals (for
CC development and screening of therapeutic agents), also for in vivo or
CC ex vivo gene therapy.
XX
SQ Sequence 300 AA;

Query Match 100.0%; Score 1634; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 3.8e-122;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRALEGPGLSLLCLVLALPVPVAVRGVAETPTYPWRDAETGERLVCAQCPPGTFVQR 60
DB 1 mralegpglsllclvlalpvpavrgvaetpypwrdaetgerlvcaqcpptfvqr 60
QY 61 PCRRDSPPTCGCPPRHRYTQFWNYLERCRYCNVLCGEREEERACHATHNRCRCRTGFF 120
DB 61 pcrrdspttcgpcprhytqfwnylercrycnvlgereeeearachathnrcrcrtgff 120
QY 121 AHAGFCLEHASCPCGAGVIAPCTPSQNTQCPCPGTFSASSSSSEOCQPHNCTALGLA 180
DB 121 ahagfclehascpagviapctpsqntqcpcpgtfsassssseqcqhpnctalgla 180
QY 181 LNVPGSSSHDTLCTSGTGFPLSTRVPGAECEERAVIDFVAFODISIKRLQLLQALEAPE 240
DB 181 lnpvgssshdtlctsgtgfplstrvpgaeceeravidfvafodisikrlqlqaleape 240
QY 241 GWGPTPRAGRAALQLKLRRLTELGAQDQALLVRLQALVARMPGLSRSVRFLPVH 300
DB 241 gwgptpragraalqlklrrrltelgagdgallvrlqalvarmpglersvrflpvh 300

RESULT 7
W97749
ID W97749 standard; Protein; 300 AA.
XX
AC W97749;
XX
DT 21-MAY-1999 (first entry)
XX
DE Human tumour necrosis factor receptor ZTNFR-5.
XX
KW ZTNFR-5; tumour necrosis factor receptor; TNFR; human;
KW cell maturation; bone cell regulation.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..23
FT Protein /note= "signal peptide"
FT Protein 24..300
FT Protein /note= "mature protein"
FT Domain 24..194
FT Region /note= "extracellular domain"
FT Region 49..71
FT Region /note= "cysteine-rich pseudo-repeat 1"
FT Region 72..113
FT Region /note= "cysteine-rich pseudo-repeat 1"
FT Region 114..151
FT Region /note= "cysteine-rich pseudo-repeat 1"
FT Region 152..194
FT Region /note= "cysteine-rich pseudo-repeat 1"
XX
XX W09904001-AL.
PN
XX
PD 28-JAN-1999.
XX

PF 21-JUL-1998; 98WO-US15072.
XX
PR 21-JUL-1997; 97US-0053203.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Farrah TM;
XX
DR WPI: 1999-132245/11.
DR N-PSDB; X07226.
XX
PT Novel tumour necrosis factor receptor ZTNFR5 - useful for
PT regulating maturation of TNF-ligand bearing cells
XX
PS Claim 1; Page 84-85; 109pp; English.
XX
CC This polypeptide comprises a new, secreted tumour necrosis factor
CC receptor (see W97749), designated ZTNFR-5. Novel ZTNFR-5 encoding
CC polynucleotides and polypeptides were initially identified by
CC querying an expressed sequence tag (EST) database for sequences
CC homologous to conserved motifs within the TNF receptor family.
CC Based on this search, a contig of 16 ESTs (see X07226) was
CC constructed. ZTNFR-5 polypeptides comprise 4 cysteine-rich repeats
CC (see also W97750-55) that are homologous to other TNF receptors, in
CC particular the soluble, secreted TNF receptor osteoprotegerin.
CC ZTNFR-5 polypeptide can be prepared by recombinant methods. The
CC polypeptide, especially the extracellular domain, can be used to
CC generate a soluble variant of ZTNFR-5. The polypeptides and
CC nucleic acids can be used to screen for ligands, agonists and
CC antagonists of ZTNFR-5. The polypeptides can be used in bone cell
CC regulation and to regulate the maturation of TNF ligand-bearing
CC cells such as T- or B-cells, lymphocytes, peripheral blood
CC mononuclear cells, polymorphonuclear leukocytes, fibroblasts or
CC haematopoietic cells.
XX
SQ Sequence 300 AA;

Query Match 100.0%; Score 1634; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 3.8e-122;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRALEGPGLSLLCLVLALPVPVAVRGVAETPTYPWRDAETGERLVCAQCPPGTFVQR 60
DB 1 mralegpglsllclvlalpvpavrgvaetpypwrdaetgerlvcaqcpptfvqr 60
QY 61 PCRRDSPPTCGCPPRHRYTQFWNYLERCRYCNVLCGEREEERACHATHNRCRCRTGFF 120
DB 61 pcrrdspttcgpcprhytqfwnylercrycnvlgereeeearachathnrcrcrtgff 120
QY 121 AHAGFCLEHASCPCGAGVIAPCTPSQNTQCPCPGTFSASSSSSEOCQPHNCTALGLA 180
DB 121 ahagfclehascpagviapctpsqntqcpcpgtfsassssseqcqhpnctalgla 180
QY 181 LNVPGSSSHDTLCTSGTGFPLSTRVPGAECEERAVIDFVAFODISIKRLQLLQALEAPE 240
DB 181 lnpvgssshdtlctsgtgfplstrvpgaeceeravidfvafodisikrlqlqaleape 240
QY 241 GWGPTPRAGRAALQLKLRRLTELGAQDQALLVRLQALVARMPGLSRSVRFLPVH 300
DB 241 gwgptpragraalqlklrrrltelgagdgallvrlqalvarmpglersvrflpvh 300
RESULT 8
W95082
ID W95082 standard; Protein; 300 AA.
XX
AC W95082;
XX
DT 20-MAY-1999 (first entry)
XX
DE Orphan receptor (HUMAN NTR-1) polypeptide.
XX

KW HUMAN NTR-1; orphan receptor; osteoprotegerin; OPG; TNFR; human;
KW tumour necrosis factor receptor; muscle disorder; bone mass; screening;
KW muscle metabolism; binding agent; cognate ligand.
XX Homo sapiens.
XX WO9907738-A2.
XX 18-FEB-1999.
XX 04-AUG-1998; 98WO-US16202.
XX 06-AUG-1997; 97US-0054869.
XX (PROC) PROCTER & GAMBLE CO.
XX (REGE-) REGENERON PHARM INC.
XX Masiaowski PJ, Morris J, Valenzuela DM;
XX WPI; 1999-167365/14.
XX N-PSDB; X22300.
XX Novel orphan human receptor polypeptide and nucleic acid - useful as
XX diagnostic reagents and for treatment of muscle disorders
XX Claim 7; Page 21; 23pp; English.
XX This represents a HUMAN NTR-1 polypeptide, a novel orphan receptor. The
XX protein is related to osteoprotegerin (OPG) and to tumour necrosis factor
XX receptor (TNFR). Host cells transformed with a vector comprising the
XX HUMAN NTR-1 nucleic acid are used for the recombinant expression of the
XX protein. HUMAN NTR-1 proteins and antibodies immuno specific for the
XX protein are useful for diagnosis and treatment of humans and animals,
XX especially muscle disorders, as the receptor is involved in regulation of
XX bone mass and muscle metabolism. HUMAN NTR-1 receptors are also useful
XX for screening for novel binding agents, and cognate ligands, which may be
XX used to treat disorders associated with HUMAN NTR-1 imbalance.

SQ Sequence 300 AA;
Query Match 100.0%; Score 1634; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 3.8e-122;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRALSGPLSLCLVLPALLPVPARGVAFETPTYPWRDAETGERLVCAQCPTGTFVQR 60
DB 1 mralepgglsllclvlpallvpavrgvafetptypwrdaetgerlvcaqcpptgftvqr 60
QY 61 PCRDRSPTTCGPPRHYTQFWNYLERCYCNVLCGEREEERACHATHNRACRRTGFF 120
DB 61 pcrdrsppttcgpprhytqfwnylercycnvlcgerereearachathnrcrtrtff 120
QY 121 AHAGFCLHASCPPGAGVYATPQTSQNTQCPCPPTGTSASSSSSEQCPHRNCTALGLA 180
DB 121 ahagfclhascpagvylatpqtsgntqcpcptgtsassssseqcpnrnctalgla 180
QY 181 LNVPGSSSHDTLCTGTFPLTRVPGAECERAVIDFVAFODISIKRLQLLQALEAPE 240
DB 181 lnvpgssshdtlctgtpfltrvpgaeceravidfvafoDISIKRLQLLQALEAPE 240
QY 241 GWGPTPRAGRAALQLKLRRLTELGAODGALLVRLQLALVARMPGLERSVREFLPVH 300
DB 241 gwptpragraaalqlklrrrltelgaodgallvrlqlalvarmpglersvrerflpvh 300

RESULT 9
Y77458
ID Y77458 standard; Protein; 300 AA.
XX
AC Y77458;
XX
DT 05-JUN-2000 (first entry)

XX Human TNF receptor-like protein, HDTEA84.
DE
XX TNF receptor family; tumour necrosis factor; HDTEA84; HSLJD378;
KW Rank-like protein; RANKL; immune disorder; inflammation; allergy;
KW immunosuppressant; antiarthritic; antirheumatoid; antiinflammatory;
KW dermatological; antithyroid.
XX
OS Homo sapiens.
XX WO200001817-A2.
PN
XX 13-JAN-2000.
PD
XX 06-JUL-1999; 99WO-US12366.
PF
XX 06-JUL-1998; 98US-0110938.
PR 13-JUL-1998; 98US-0114466.
PR 23-JUL-1998; 98US-0093897.
PR 12-AUG-1998; 98US-0132968.
PR 18-AUG-1998; 98US-0136214.
PR 11-SEP-1998; 98US-0099999.
XX
XX (SCHE) SCHERING CORP.
PA
XX Bates EEM, Lebecque SJE, Murphy EE, Mattson JD, Gorman DM;
PI Hedrick JA, Wang L, Zlotnik A, Murgolo NJ, Greene JR, Johnston JA;
PI Bazan JF, Mahony D, Lees EM;
XX WPI; 2000-171015/15.
DR N-PSDB; Z92404.
DR
XX New isolated mammalian genes, used to develop products for treating
XX e.g. immune, inflammatory or allergic abnormalities, cancers or
XX degenerative conditions
XX Claim 24; Page 157; 218pp; English.
XX The invention relates to a number of primate and/or rodent proteins, and
XX the genes which encode them. The invention encompasses human dendritic
XX cell prostaglandin transporter (DC-PGT); the TNF (tumour necrosis
XX factor) receptor family-related proteins HDTEA84, HSLJD378 and RANKL;
XX human CC chemokine HCC5; human deubiquitinating proteins Dub1 and Dub
XX 12; human MD-1 and human and murine MD-2 proteins, which exhibit the
XX properties of ligands for proteins comprising a leucine-rich motif (LRR);
XX human cyclin E2; cDNAs encoding these proteins; and antibodies against
XX these proteins. The proteins can be used for modulating the physiology or
XX development of a cell. They can be used to mediate uptake of substrates
XX (e.g., prostaglandin-like molecules), to modulate or mediate cellular
XX interactions (e.g., induce or prevent trafficking, proliferation, or
XX differentiation of cells), or are intracellular proteins which are
XX important in various cellular processes such as the deubiquitination of
XX proteins or cell cycle regulation. The products can be used for
XX treating medical conditions such as immune, inflammatory or allergic
XX disorders, or abnormal cellular proliferation, for example, cancers or
XX degenerative conditions. They can be used to modulate immune responses in
XX disease states e.g., autoimmune disorders, including rheumatoid
XX arthritis, systemic lupus erythematosus, Hashimoto's autoimmune
XX thyroiditis, as well as acute and chronic inflammatory responses in which
XX T cell activation, expansion, and/or immunological T cell memory play an
XX important role. Sequences Y77458-Y77461 and Y77463-Y77468 represent
XX TNF receptor family-related proteins. Y77458 is the human protein
XX HDTEA84. Y77459-Y77461 are human HSLJD378 proteins, Y77463 is
XX murine Rank-like protein RANKL, and Y77466-Y77468 are human RANKL
XX proteins.
XX Sequence 300 AA;
SQ

Query Match 99.1%; Score 1620; DB 21; Length 300;
Best Local Similarity 99.3%; Pred. No. 4.9e-121;
Matches 298; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MRALEGPGSLLCVLALPALLPVPVAVGVAETPTYPWRDAETGERLVCAQCPTGTFVQR 60
 Db 1 mralepgpalsllclvlalpallpvpavrgvaetptypwrdaetgerlvcaqcpptgfvr 60
 Qy 61 PCRDSPTTCGCPPRHYTQFWNYLERCRYCNVLCGEREEARACHATHNACRCRTGFF 120
 Db 61 pccrdspttcgcpprhytqfwnylercrycnvlgereearachathnacrctgff 120
 Qy 121 AHAGFCLHASCPCGAGVIAPGTPSONTCOCPPGTFSSSSSQCOCPHRNCTALGLA 180
 Db 121 ahagfclhascpagviaptpsqntcqcpcppgtfsasssseqcqhnrctalgia 180
 Qy 181 LNVPGSSSHDITCTSGTGFPLSTRVPAEGERAVIDFVAFQDISIKRLQLLQALEAPE 240
 Db 181 lnvpgssshdtlctsgtgpflstrvpaeeeceravidfvafqdisikrlqlqaleape 240
 Qy 241 GWGPTPRAGRAALQIKLRRRLTELIGAODGALLVRLQLALRVARMPGLERSVRERFLPVH 300
 Db 241 gwgptpragraaglkilrrrtelligadgallvrlqlalrvarmpglersvrerflpvh 300

RESULT 10

Y42183
 ID Y42183 standard; Protein; 302 AA.

XX AC Y42183;

XX DT 17-DEC-1999 (first entry)

XX DE Human FLINT #2 protein sequence.

XX KW Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
 KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
 KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
 KW reperfusion-associated injury; aplastic anaemia; differentiation;
 KW growth; myelodysplastic syndrome; pancytopenic condition;
 KW myocardial ischaemia.

XX OS Homo sapiens.

XX PN W09950413-A2.

XX PD 07-OCT-1999.

XX PE 30-MAR-1999; 99WO-US06797.

XX PR 30-MAR-1998; 98US-0079856.

XX PR 20-MAY-1998; 98US-0086074.

XX PR 09-SEP-1998; 98US-0099643.

XX PR 17-DEC-1998; 98US-0112577.

XX PR 18-DEC-1998; 98US-0112703.

XX PR 18-DEC-1998; 98US-0112933.

XX PR 22-DEC-1998; 98US-0113407.

XX PA (ELIL) LILLY & CO ELI.

XX PI Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;

XX PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;

XX PI Song HY, Wang J, Wu X, Zuckerman SH;

XX DR WPI; 1999-591319/50.

XX DR N-PSDB; 225376.

XX PT Use of mature FLINT for treating acute liver failure, inflammation,

XX PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic

XX PT and proinflammatory activity

XX PS Example 2; Fig 2; 99pp; English.

CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
 CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
 CC injury or disorder such as hypercoagulation (including use with
 CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
 CC or disorder, Type I diabetes, cancer, cell damage or damage to an
 CC innocent bystander tissue that is induced by a chemotherapeutic agent or
 CC therapeutic irradiation, treating haematopoietic progenitor cells that
 CC have been exposed to therapeutic radiation or chemotherapy, aplastic
 CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
 CC also used for promoting the growth or differentiation of a haematopoietic
 CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
 CC resulting from abnormal myocardial ischaemia. The present sequence
 CC represents human FLINT.

XX SQ Sequence 302 AA;

Query Match 98.5%; Score 1610; DB 20; Length 302;

Best Local Similarity 98.7%; Pred. No. 3.1e-120;

Matches 298; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

Qy 1 MRALEGPGSLLCVLALPALLPVPVAVGVAETPTYPWRDAETGERLVCAQCPTGTFVQR 60
 Db 1 mralepgpalsllclvlalpallpvpavrgvaetptypwrdaetgerlvcaqcpptgfvr 60
 Qy 61 PCRDSPTTCGCPPRHYTQFWNYLERCRYCNVLCGEREEARACHATHNRA--CRCRTG 118
 Db 61 pccrdspttcgcpprhytqfwnylercrycnvlgereearachathnacrctg 120
 Qy 119 FFAHAGFCLHASCPCGAGVIAPGTPSONTCOCPPGTFSSSSSQCOCPHRNCTALG 178
 Db 121 ffahagfclhascpagviaptpsqntcqcpcppgtfsasssseqcqhnrctalg 180
 Qy 179 LALNVPGSSSHDITCTSGTGFPLSTRVPAEGERAVIDFVAFQDISIKRLQLLQALEA 238
 Db 181 lalnvpgssshdtlctsgtgpflstrvpaeeeceravidfvafqdisikrlqlqalea 240
 Qy 239 PEGWGPTPRAGRAALQIKLRRRLTELIGAODGALLVRLQLALRVARMPGLERSVRERFLP 298
 Db 241 pegwaptpragraaglkilrrrtelligadgallvrlqlalrvarmpglersvrerflp 300
 Qy 299 VH 300
 Db 301 wh 302

RESULT 11

Y42184

ID Y42184 standard; Protein; 271 AA.

XX AC Y42184;

XX DT 17-DEC-1999 (first entry)

XX DE Human mFLINT #1 protein sequence.

XX KW Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
 KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
 KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
 KW reperfusion-associated injury; aplastic anaemia; differentiation;
 KW growth; myelodysplastic syndrome; pancytopenic condition;
 KW myocardial ischaemia.

XX OS Homo sapiens.

XX PN W09950413-A2.

XX PD 07-OCT-1999.

XX PF 30-MAR-1999; 99WO-US06797.

XX PR 30-MAR-1998; 98US-0079856.

XX PR 20-MAY-1998; 98US-0086074.

CC The present invention describes therapeutic applications of mature FLINT
 CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
 CC (mFLINT), which is a member of the tumour necrosis factor receptor
 CC superfamily, is used for treating acute liver failure, inflammation of

```

PR 09-SEP-1998; 98US-0099643.
PR 17-DEC-1998; 98US-0112577.
PR 18-DEC-1998; 98US-0112703.
PR 18-DEC-1998; 98US-0112933.
PR 22-DEC-1998; 98US-0113407.
XX
PA (ELIL ) LILLY & CO ELI.
XX
XX Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
PI Song HY, Wang J, Wu X, Zuckerman SH;
XX
XX WPI; 1999-591319/50.
DR N-PSDB; Z25377.
XX
XX use of mature FLINT for treating acute liver failure, inflammation,
PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
PT and proinflammatory activity
XX
XX Claim 31; Fig 3; 99pp; English.
PS
XX
XX The present invention describes therapeutic applications of mature FLINT
XX (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
XX (mFLINT), which is a member of the tumour necrosis factor receptor
XX superfamily, is used for treating acute liver failure, inflammation of
XX the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
XX with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
XX injury or disorder such as hypercoagulation (including use with
XX thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
XX or disorder, Type I diabetes, cancer, cell damage or damage to an
XX innocent bystander tissue that is induced by a chemotherapeutic agent or
XX therapeutic irradiation, treating haematopoietic progenitor cells that
XX have been exposed to therapeutic radiation or chemotherapy, aplastic
XX anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
XX also used for promoting the growth or differentiation of a haematopoietic
XX progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
XX resulting from abnormal myocardial ischaemia. The present sequence
XX represents human mFLINT.
XX
XX Sequence 271 AA;

Query Match 91.2%; Score 1491; DB 20; Length 271;
Best Local Similarity 100.0%; Pred. NO. 7.3e-111;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 VAETPTYPWRDAETGERLVCACQCPGTFVQPCRRDSPTTCGCPGPRHYTQFWNYLERCR 89
Db 1 vaetptyprdaetgerlvcaqcpptfvqpcrrdspttcgpcpprhytqfwnlylerc 60

QY 90 YCNVLGGEREEARACHATHNRA--CRCTRGTFFAHAGFCLEHASCPPGAGVIAPGTPSQNTQ 149
Db 1 ycnvlcgereearachathnra--crctrgtffahagfclehascpaggviapgtpsqntq 120

QY 150 CQPCPPGTFSSSSSEQCQPHNCTALGLALNVPSSSHDTLCTSCGTGFLSTRVPGAE 209
Db 121 cpcppgtfssssseqcqphnctaqlalnvpssshdtlctscgtgflstrvpgae 180

QY 210 ECERAVIDFVAQDISIKELRLQALAEPEGWGPTPRAGRAALQKLRRRLTELLGAQD 269
Db 181 eceravidfvaqdisikrlqalaepegwgptpragraalqlkrrrltellgaqd 240

QY 270 GALLVRLQALRVARMPGLSVRERFLPVH 300
Db 241 gallvrlqalrvamrglsvrerflpvh 271

RESULT 12
Y42185
ID Y42185 standard; protein: 273 AA.
XX
AC Y42185;
XX
```

```

DT 17-DEC-1999 (first entry)
XX
XX Human mFLINT #2 protein sequence.
XX
XX Human; FLINT; mFLINT; OPC3; tumour necrosis factor receptor; FasL;
KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
KW reperfusion-associated injury; aplastic anaemia; differentiation;
KW growth; myelodysplastic syndrome; pancytopenic condition;
KW myocardial ischaemia.
XX
XX Homo sapiens.
XX OS
XX WO9950413-A2.
XX
XX 07-OCT-1999.
PD
XX
XX 30-MAR-1999; 99WO-US06797.
PF
XX
XX 30-MAR-1998; 98US-0079856.
PR
XX 20-MAY-1998; 98US-0086074.
PR
XX 09-SEP-1998; 98US-0099643.
PR
XX 17-DEC-1998; 98US-0112577.
PR
XX 18-DEC-1998; 98US-0112703.
PR
XX 18-DEC-1998; 98US-0112933.
PR
XX 22-DEC-1998; 98US-0113407.
XX
XX (ELIL ) LILLY & CO ELI.
XX
XX Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
PI Song HY, Wang J, Wu X, Zuckerman SH;
XX
XX WPI; 1999-591319/50.
DR N-PSDB; Z25378.
XX
XX Use of mature FLINT for treating acute liver failure, inflammation,
XX cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
XX and proinflammatory activity
XX
XX Example 2; Fig 4; 99pp; English.
XX
XX The present invention describes therapeutic applications of mature FLINT
XX (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
XX (mFLINT), which is a member of the tumour necrosis factor receptor
XX superfamily, is used for treating acute liver failure, inflammation of
XX the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
XX with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
XX injury or disorder such as hypercoagulation (including use with
XX thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
XX or disorder, Type I diabetes, cancer, cell damage or damage to an
XX innocent bystander tissue that is induced by a chemotherapeutic agent or
XX therapeutic irradiation, treating haematopoietic progenitor cells that
XX have been exposed to therapeutic radiation or chemotherapy, aplastic
XX anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
XX also used for promoting the growth or differentiation of a haematopoietic
XX progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
XX resulting from abnormal myocardial ischaemia. The present sequence
XX represents human mFLINT.
XX
XX Sequence 273 AA;

Query Match 89.8%; Score 1467; DB 20; Length 273;
Best Local Similarity 98.5%; Pred. NO. 5.8e-109;
Matches 269; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 30 VAETPTYPWRDAETGERLVCACQCPGTFVQPCRRDSPTTCGCPGPRHYTQFWNYLERCR 89
Db 1 vaetptyprdaetgerlvcaqcpptfvqpcrrdspttcgpcpprhytqfwnlylerc 60

QY 90 YCNVLGGEREEARACHATHNRA--CRCTRGTFFAHAGFCLEHASCPPGAGVIAPGTPSQN 147
Db 1 ycnvlcgereearachathnra--crctrgtffahagfclehascpaggviapgtpsqn 147
```

Db 61 ycnvlgereearachathnracrcrtgtffahagfclchascppgagviapgtpsqn 120
Qy 148 TQCCPPPTFSASSSSQEQPHNCTALGLALNVPSSSHDITLCTSGTFPLSTRVPG 207
Db 121 tqccppptfsassseqqphnctaiglanvpgsshdltctscgtgfplstrvpg 180
Qy 208 AEEGERAVIDFVAFQDISIKRLQLQALEAPEGWGTPPRAGRAALQLKRRRLTELLGA 267
Db 181 aeceravidfvaqdisikrlqlqaleapegwaptpragraalqlkrrrltellga 240
Qy 268 QDGALLVRLQALRVARMPGLERSVRERFLPVH 300
Db 241 qdgallvrlqalrvarmpglersvrerflpwh 273

RESULT 13
Y28449
ID Y28449 standard; Protein; 245 AA.

AC Y28449;
XX
XX
DT 29-SEP-1999 (first entry)
XX

A human tumour necrosis factor-R2-like proteins (TR2P)-1.

Human tumour necrosis factor-R2-like protein; TR2p; achondroplasia;
osteoporosis; developmental disorder; Cushing's syndrome;
muscular dystrophy; epilepsy; hereditary neuropathy;
Charcot-Marie-Tooth disease; neurofibromatosis; hypothyroidism;
hydrocephalus; seizure disorder; cerebral palsy; spinal bifida;
congenital glaucoma; cataract; sensorineural hearing loss;
reproductive disorder; infertility; ovulatory defect; endometriosis;
autoimmune disorder; ectopic pregnancy; teratogenesis; spermatogenesis;
immunological disorder; AIDS; Addison's disease; allergy; bronchitis;
atherosclerosis; diabetes mellitus; Chron's disease; lupus;
irritable bowel syndrome; multiple sclerosis; infection;
neoplastic disorder; adenocarcinoma; leukaemia; lymphoma; melanoma;
myeloma; sarcoma.

XX Homo sapiens.

XX W09931128-A2.

XX 24-JUN-1999.

XX 02-DEC-1998; 98WO-US25649.

XX 16-DEC-1997; 97US-0991945.

XX (INCY-) INCYTE PHARM INC.

XX Au-Young J, Bandman O, Hillman JL, Kaser MR, Tang YT;

XX WPI; 1999-457916/38.

XX N-PSDB; X89503.

XX New tumour necrosis factor-R2-like protein - useful in the treatment
of osteogenesis, developmental, reproductive, immunological and
neoplastic disorders

XX Claim 1; Fig 1A-C; 81pp; English.

XX The present sequence represents a human tumour necrosis factor-R2-like
protein (TR2P). The protein is used to treat and prevent osteogenesis,
developmental, reproductive, immunological and neoplastic disorders, and
also to diagnose disorders associated with TR2 protein expression. Such
disorders include osteogenesis disorders such as achondroplasia and
osteoporosis, developmental disorders such as Cushing's syndrome,
muscular dystrophy, and epilepsy, hereditary neuropathies such as
Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism,
hydrocephalus, seizure disorders such as cerebral palsy and spinal
bifida, congenital glaucoma, cataract, or sensorineural hearing loss,
reproductive disorders such as infertility, ovulatory defects and

CC endometriosis, autoimmune disorders, ectopic pregnancy and teratogenesis,
disruption of spermatogenesis, immunologic disorders such as AIDS,
CC Addison's disease, allergies, bronchitis, atherosclerosis, diabetes,
CC mellitus, Chron's disease, lupus and irritable bowel syndrome, multiple
CC sclerosis, viral, fungal, helminthic, parasitic and protozoal infections,
CC and neoplastic disorders including adenocarcinoma, leukaemia, lymphoma,
CC melanoma, myeloma, sarcoma, and teratocarcinoma.
XX
SQ Sequence 245 AA;

Query Match 83.4%; Score 1362; DB 20; Length 245;
Best Local Similarity 99.6%; Pred. No. 1.1e-100;
Matches 244; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MRALEGFGLSLLLCLVLPALLPVPVAVRGVAETPTYPWRDAETGERLVCAQCPGTFVQR 60
Db 1 mralegpglsllclvlpalpvpavrgvaetptypwrdaetgerlvcaqcpptfvqr 60
XX

Qy 61 PCRRDSPPTTCGPPRHYTQFNWYLERCRYCNVLCGEREEARACHATHNACRCRTGFF 120
Db 61 pcrrdsppttcgpprhytqfnwylerycnvlgereearachathnacrctgff 120
XX

Qy 121 AHAGFCLEHASCPCGAGVIAPGTPSQNTQCPCPPGTFSSSSSEQCQPHNCTALGLA 180
Db 121 ahagfclchascppgagviapgtpsqntqcpcppgtfssssseqcqphnctalgla 180
XX

Qy 181 LNVGSSSHDITLCTSGTFPLSTRVPGAEGERAVIDFVAFQDISIKRLQLQALEAPE 240
Db 181 lnvpgssshdltctscgtgfplstrvpgaeeceravidfvaqdisikrlqlqaleape 240
XX

Qy 241 GWGPT 245
Db 241 dwgpt 245
XX

RESULT 14

W93585

ID W93585 standard; Protein; 215 AA.

XX W93585;

XX 18-JUN-1999 (first entry)

XX Human hAPO6 protein.

XX Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
developmental abnormality; gestational abnormality; prostate cancer;
KW APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
KW apoptosis; human.
XX

OS Homo sapiens.

XX W09911791-A2.

XX 11-MAR-1999.

XX 04-SEP-1998; 98WO-US18393.

XX 05-SEP-1997; 97US-0924634.

XX (UNIW) UNIV WASHINGTON.

XX Chaudhary PM;

XX WPI; 1999-205191/17.

XX N-PSDB; X23419.

XX New Tumor Necrosis Factor family receptor polypeptides and ligands -
PT useful for diagnosis and treatment of prostate cancer and
PT developmental or gestational abnormalities
XX

PS Claim 29; Fig 9; 156pp; English.

XX This invention describes isolated Tumor Necrosis Factor (TNF) family

CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active

CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or

CC their active fragments. APO4 is useful for diagnosing prostate cancer

CC by determining levels of APO4 in an individual. Prostate cancer can also

CC be treated using APO4 selective binding agents linked to a therapeutic

CC moiety. APO4 polypeptides are also useful for identifying selective

CC binding agents, useful in diagnosis/treatment of disease by binding of

CC agents to the polypeptide/active fragment which is extracellular, or

CC expressed on the cell surface. The binding is preferably performed in

CC vivo. APO4 polypeptides/active fragments are also useful for screening

CC for agonists and antagonists by binding and observing the change in APO4

CC activity. Effective pharmacological agents useful in diagnosis or

CC treatment of disease are also identified using APO4 polypeptides/active

CC fragments and APO4 signal transducer molecules that specifically interact

CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4

CC activity. The method is performed in vivo or in vitro. APO polypeptides

CC are all useful as immunogens for preparing antibodies. APO4 is also

CC useful for diagnosis/treatment of developmental or gestational

CC abnormalities. APO8 was transfected to human breast carcinoma cell line

CC MCF-7, and induced apoptosis.

XX Sequence 215 AA;

SQ

Query Match 70.6%; Score 1153; DB 20; Length 215;

Best Local Similarity 100.0%; Pred. No. 3.3e-84;

Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 ERCRYCNVLCGEREEARACHATHNRACRGTGFFAHAGFCLHASCPCPGAGVIAPGTPS 145

DB 1 ERCRYCNVLCGEREEARACHATHNRACRGTGFFAHAGFCLHASCPCPGAGVIAPGTPS 60

QY 146 QNTQCQPCPGPTGFSASSSSSEQCQPHRNCRTALGLALNVPGSSSHDTLCTCTGFPPLSTRV 205

DB 61 QNTQCQPCPGPTGFSASSSSSEQCQPHRNCRTALGLALNVPGSSSHDTLCTCTGFPPLSTRV 120

QY 206 PGAECEERAVIDFVAFQDISIKRLQLQALAEPEGWGPTPRAGRAALQLKRLRLTELL 265

DB 121 PGAECEERAVIDFVAFQDISIKRLQLQALAEPEGWGPTPRAGRAALQLKRLRLTELL 180

QY 266 GAQDGLLVRLQLALRVAMPGLERSVRERFLPVH 300

DB 181 GAQDGLLVRLQLALRVAMPGLERSVRERFLPVH 215

RESULT 15

ID Y22222 standard; Protein; 153 AA.

XX Y22222;

XX AC

XX DT

XX 16-SEP-1999 (first entry)

DE Human TNFR superfamily soluble receptor protein sequence.

XX TNFL1; human; TNFR superfamily; tumour necrosis factor ligand; TNF;

KW tumour necrosis factor receptor; TNFR superfamily; cell proliferation;

KW cell differentiation; cytokine production; immunoglobulin; hyperplasia;

KW apoptosis inducer; activated T cell; autoimmune disease; inhibitor;

KW myasthenia gravis; insulin-dependent diabetes mellitus; endotoxin shock;

KW rheumatoid arthritis; multiple sclerosis; systemic lupus erythematosus;

KW tumour; proliferative disorder; neoplasia; dysplasia; immunocompetence;

KW lymphoid organogenesis; bacterial resistance; contact hypersensitivity;

XX delayed type sensitivity; therapy.

OS Homo sapiens.

XX

PN WC0933980-A2.

XX

PD 08-JUL-1999.

XX PF 22-DEC-1998; 98WO-US27474.

XX PR 16-DEC-1998; 98US-0212270.

XX PR 30-DEC-1997; 97US-0068959.

XX PA (CHIR) CHIRON CORP.

XX PI Kassam A, Lamson G, Pot D, Tribouley C;

XX WPI; 1999-405508/34.

XX DR N-ESDB; X84621.

XX New tumour necrosis factor ligands, useful for induction of cell

XX death and/or proliferation of cells

XX Claim 1; Page 61; 69pp; English.

XX This sequence represents a tumour necrosis factor receptor (TNFR)

CC superfamily soluble protein of the invention. The invention also relates

CC to tumour necrosis factor (TNF) ligand (TNFL) family proteins. The TNFL

CC proteins play regulatory roles in cell proliferation and/or

CC differentiation, e.g. they can induce production of cytokines,

CC immunoglobulins, etc. A variety of diseases can be treated by modulating

CC the activity of TNFL proteins, e.g. they can induce apoptosis of

CC activated T cells but rescue resting T cell from apoptosis. TNFL

CC polypeptides can therefore be used to treat autoimmune diseases, such as

CC myasthenia gravis, insulin-dependent diabetes mellitus, rheumatoid

CC arthritis, multiple sclerosis, and systemic lupus erythematosus. TNFL

CC proteins also have tumour stimulating properties, so tumours can be

CC treated by inhibiting the expression or activity of TNFL. Other

CC proliferative disorders, such as neoplasias, dysplasias, and hyperplasia

CC can also be treated using TNFL inhibitors. The TNFL polypeptides and

CC polynucleotides can also be used to enhance or decrease TNF activity,

CC thus providing therapeutic benefits such as induction of cell death,

CC lymphoid organogenesis, or host bacterial resistance, and inhibition of

CC endotoxin shock, contact hypersensitivity, delayed type sensitivity or

CC immunocompetence of a transplant recipient. TNF and its receptors play a

CC major role in host defence and immunosurveillance. As such, there is a

CC need to identify new members of TNFR families. This invention provides

CC this need.

XX Sequence 153 AA;

SQ

Query Match 51.5%; Score 841; DB 20; Length 153;

Best Local Similarity 100.0%; Pred. No. 1.2e-59;

Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 LERCRCNVLCGEREEARACHATHNRACRGTGFFAHAGFCLHASCPCPGAGVIAPGTP 144

DB 1 LERCRCNVLCGEREEARACHATHNRACRGTGFFAHAGFCLHASCPCPGAGVIAPGTP 60

QY 145 SONTQCQPCPGPTGFSASSSSSEQCQPHRNCRTALGLALNVPGSSSHDTLCTCTGFPPLSTR 204

DB 61 SONTQCQPCPGPTGFSASSSSSEQCQPHRNCRTALGLALNVPGSSSHDTLCTCTGFPPLSTR 120

QY 205 VPGAEECEERAVIDFVAFQDISIKRLQLQALAE 237

DB 121 VPGAEECEERAVIDFVAFQDISIKRLQLQALAE 153

Search completed: January 30, 2001, 16:45:18

Job time: 578 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:38:30 ; Search time 149.64 seconds
(without alignments)
136.128 Million cell updates/sec

Title: US-09-518-931-2
Perfect score: 1634
Sequence: 1 MRALGPGLSLLCLVALPA.....RVARMPGLSVRRERFLPVH 300

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_66: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	351.5	21.5	461	1 A35356	tumor necrosis fac
2	333.5	20.4	459	2 I48894	gene murine tumour
3	332.5	20.3	474	2 B38634	tumor necrosis fac
4	315	19.3	435	2 I54182	tumor necrosis fac
5	262.5	16.1	349	2 D72175	G2R protein - vari
6	262.5	16.1	349	2 D36858	gene G4R protein -
7	262	16.0	348	2 T28623	hypothetical prote
8	236.5	14.5	325	2 B43692	T2 protein - rabbi
9	226	13.8	277	2 I37552	OX40 homolog - hum
10	215	13.2	326	1 GQVZML	T2 protein - myxom
11	214	13.1	271	2 S12783	OX40 antigen precu
12	211	12.9	277	2 A60771	B-cell activation
13	203	12.4	289	2 A46515	B cell-associated
14	203	12.4	305	2 A46476	B cell-associated
15	198.5	12.1	272	2 I48700	gene ox40 protein
16	186.5	11.4	595	2 A42086	CD30 antigen precu
17	185	11.3	256	2 B32393	T-cell antigen 4-1
18	176	10.8	416	1 JN0006	nerve growth facto
19	175.5	10.7	427	1 GQHUN	nerve growth facto
20	174	10.6	255	2 I38426	4-LBB - human
21	173	10.6	255	2 J70752	lymphocyte activat
22	170	10.4	425	1 A26431	nerve growth facto
23	159.5	9.8	260	1 A46517	CD27 antigen precu
24	155.5	9.5	327	2 A46484	apoptosis-mediati
25	148.5	9.1	1574	2 T13954	MEGF6 protein - ra
26	148	9.1	250	1 A49033	CD27 antigen precu
27	147.5	9.0	5376	2 T42215	zonadhesin - mouse
28	145	8.9	335	2 A40036	apoptosis-mediati
29	144	8.8	324	2 JC2395	Fas antigen precu

ALIGNMENTS

RESULT 1

A35356

tumor necrosis factor receptor type 2 precursor - human
N:Alternate names: 75K tumor necrosis factor receptor
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence.revision 10-Sep-1999 #text.change 10-Sep-1999
C:Accession: A35356; A36475; A48416; A36007; A23666; B35010; I38094
R:Smith, C.A.; Davis, T.; Anderson, D.; Solam, L.; Beckmann, M.P.; Jerzy, R.; Dower, Science 248, 1019-1023, 1990
A:Title: A receptor for tumor necrosis factor defines an unusual family of cellular a
A:Reference number: A35356; MUID:90260639
A:Accession: A35356
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-461 <SMI>
A:Cross-references: GB:M32315; NID:gl89185; PIDN:AAA5929.1; PID:gl89186
R:Kohn, T.; Brewer, M.T.; Baker, S.L.; Schwartz, P.E.; King, M.W.; Hale, K.K.; Squir Proc. Natl. Acad. Sci. U.S.A. 87, 8331-8335, 1990
A:Title: A second tumor necrosis factor receptor gene product can shed a naturally oc
A:Reference number: A36475; MUID:91045991
A:Accession: A36475
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-195, 'R', 197-461 <KOH>
A:Cross-references: GB:M55994; GB:M38549; NID:g339757; PIDN:AAA36755.1; PID:g339758, R:Dembic, Z.; Loetscher, H.; Gubler, U.; Pan, Y.C.; Lahm, H.W.; Gentz, R.; Brockhaus, Cytokine 2, 231-237, 1990
A:Title: Two human TNF receptors have similar extracellular, but distinct intracellular
A:Reference number: A48416; MUID:91370690
A:Accession: A48416
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 23-461 <DEM>
A:Cross-references: GB:S63368; NID:g235648; PIDN:AAB19824.1; PID:g235649
A:Note: sequence extracted from NCBI backbone (NCBIN:63368, NCBI:63371)
R:Heiler, R.A.; Song, K.; Onasch, M.A.; Fischer, W.H.; Chang, D.; Ringold, G.M. Proc. Natl. Acad. Sci. U.S.A. 87, 6151-6155, 1990
A:Title: Complementary DNA cloning of a receptor for tumor necrosis factor and demons
A:Reference number: A36007; MUID:90349572
A:Accession: A36007
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 116-140, 'P', 142-195, 'R', 197-362, 'T', 364-461 <HEL>
A:Cross-references: GB:M35857; NID:g339751; PIDN:AAA63362.1; PID:g339752
R:Loetscher, H.; Schlaeger, E.J.; Lahm, H.W.; Pan, Y.C.E.; Lesslauer, W.; Brockhaus, J. Biol. Chem. 265, 20131-20138, 1990
A:Title: Purification and partial amino acid sequence analysis of two distinct tumor
A:Reference number: A23666; MUID:91056048
A:Accession: A23666
A>Status: preliminary
A:Molecule type: protein
A:Residues: 23-40;65-69;136-141;300-306 <LOE>

furin (EC 3.4.21.7)
hypothetical prote
FAS soluble protei
tumor necrosis fac
tumor necrosis fac
notch3 protein - h
membrane glycoprot
serine proteinase
laminin alpha 5 ch
laminin gamma 2 ch
tumor necrosis fac
protein-tyrosine k
adhesive ligand ep
laminin alpha-2 ch
tumor necrosis fac
hypothetical prote

A:Residues: 1-349 <BLI>
A:Cross-references: GB:X69198; NID:g456758; PIDN:CAA49137.1; PID:g457087
A:Experimental source: strain India-1967, ssp. major, isolate Ind3
R:Kolykhalov, A.A.; Blinov, V.M.; Gytarov, V.V.; Pozdnyakov, S.G.; Chizhikov, V.E.; F
submitted to the EMBL Data Library, April 1992
A:Description: Nucleotide sequence analysis of the region of variola virus XhoI F O H
A:Reference number: S46868
A:Accession: S46888
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-349 <KOL>
A:Cross-references: EMBL:X67117; NID:g516428; PIDN:CAA47540.1; PID:g516449
A:Experimental source: strain India-1967, isolate Ind3
R:Shchelkunov, S.N.; Blinov, V.M.; Sandakhschiev, L.S.
FEBS Lett. 319, 80-83, 1993

Qy 124 GFCLHASCPPGAGVIAPGTPSQNTQCPCPPGTFSSSSSEOCQPHRNCTALGLALNV 183
Db 113 -TQPLDSYKFG-----VDCAPCPGHE--SPGDQACKPWTNCTLAGKHTLQ 156
Qy 184 PGSSSHDTLCTS---CTGFPLSTRVPGAEECERAVIDFVAFODISIKRLQLLQALEAPE 240
Db 157 PASNSSDAICEDROPDPATQPOETQGPAPPI-----TVQPT 193
Qy 241 GW-----GPTPR-----AGRAALQKLRRRLTELLGAQDGAALLVRLLOALRVARMP 286
Db 194 AMPRTSQGPSTRPVEVPGGRAVAAILGLVLGLLGPL--AILLALYLLRRDRQLP 247
RESULT 10
GOVZML
T2 protein - myxoma virus (strain Lausanne)
C:Species: myxoma virus
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 18-Jun-1999
C:Accession: A40566
R:Upton, C.; Macen, J.L.; Schreiber, M.; Mcfadden, G.
Virolgy 184, 370-382, 1991
A:Title: Myxoma virus expresses a secreted protein with homology to the tumor necrosis f
A:Reference number: A40566; MUID:91335768
A:Accession: A40566
A:Molecule type: DNA
A:Residues: 1-326 <UPT>
A:Cross-references: GB:M95181; GB:M37976; NID:g332309; PIDN:AAA46632.1; PID:g332310
C:Superfamily: myxoma virus T2 protein; NGF receptor repeat homology
C:Keywords: glycoprotein
F:64-105/Domain: NGF receptor repeat homology <NG2>
F:106-147/Domain: NGF receptor repeat homology <NG3>
F:66,181,205,238/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 13.2%; Score 215; DB 1; Length 326;
Best Local Similarity 29.3%; Pred. No. 4.4e-09;
Matches 58; Conservative 22; Mismatches 96; Indels 22; Gaps 5;
Qy 12 LCLVLALPALL-----PVPVARGVAETFTYWRDAETGERLVCACQCPGTFVQRCRRDS 66
Db 4 LTLLEYVACYVGGAPYGADRGKRGNDY-----EXDGLCTSCPPGSYASRLCGPGS 57
Qy 67 PTTGCPCPRHRYTOFWNYLERCVNVLGGEREEARACHATHNRACRCRTGFEFA---- 121
Db 58 DTVCSPCANETFTASTNAPACVSCRGCTGHLSESQSCDTRDVCOCAGNYCLLKQ 117
Qy 122 -HAGFLEHASCPPGAGVIAPGTPSQNTQCPCPPGTFSSSSSEOCQPHRNCTALGLA 180
Db 118 EGCRIKAPKTKCPAGYGV-SGHTRTGDVLCTKCPRTYSDAVSSTETCTSSFNYSVFN 176
Qy 181 LNVPGSSSHDTLCTSGT 198
Db 177 L-----YPVNDTSCRTAG 190
RESULT 11
SI2783
OX40 antigen precursor - rat
N:Alternate names: nerve growth factor receptor homolog
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 05-Nov-1999
C:Accession: SI2783; S08036
R:Mallett, S.; Fossum, S.; Barclay, A.N.
EMBO J. 9, 1063-1068, 1990
A:Title: Characterization of the MRC OX40 antigen of activated CD4 positive T lymphocyte
A:Reference number: SI2783; MUID:90214614
A:Accession: SI2783
A:Molecule type: mRNA
A:Residues: 1-271 <MAL>
A:Cross-references: EMBL:X17037; NID:g57830; PIDN:CAA34897.1; PID:g57831
C:Superfamily: CD27 antigen; NGF receptor repeat homology
C:Keywords: growth factor receptor; transmembrane protein

F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-271/Product: OX40 antigen #status predicted <MAT>
F:211-235/Domain: transmembrane #status predicted <MM>
Query Match 13.1%; Score 214; DB 2; Length 271;
Best Local Similarity 30.1%; Pred. No. 4.4e-09;
Matches 58; Conservative 23; Mismatches 64; Indels 48; Gaps 9;
Qy 10 SLLCLVLALPALLPVPVARGVAETFTYWRDAETGERLVCACQCPGTFVQRCRRDSPTT 69
Db 10 AFLLLGLSLGVTKLNCVK-----DTYP-----SGHK-CCRECQCHGMVSRCDHTRDTV 58
Qy 70 CGCPCPRHRYTOFWNY--LERCRYCNVLGGEREEARACHATHNRACRCRTGFEFAHGFCL 127
Db 59 CHPCEPGFYNEAVNYDTCKQCTQCNRSG--SELKQNTPTPTDTVCQR----- 105
Qy 128 EHACPPGAGVIAPGT-PSQNT-----QCPCPPGTFSSSSSEOCQPHRNCTALGLA 180
Db 106 -----PCTQPRQDSSHLKLGVDVCPGPHF--SPGSNOACKPWTNCTLSGKQ 150
Qy 181 LNVPGSSSHDTLCT 193
Db 151 IRRPASNSLDTVC 163
RESULT 12
A60771
B-cell activation protein CD40 precursor - human
N:Alternate names: B-cell surface antigen Bp50
C:Species: Homo sapiens (man)
C:Date: 03-Jun-1993 #sequence_revision 03-Feb-1994 #text_change 21-Jul-2000
C:Accession: S04460; A60771
R:Stamenkovic, I.; Clark, E.A.; Seed, B.
EMBO J. 8, 1403-1410, 1989
A:Title: A B-lymphocyte activation molecule related to the nerve growth factor recept
A:Reference number: S04460; MUID:89356608
A:Accession: S04460
A:Molecule type: mRNA
A:Residues: 1-277 <STA>
A:Cross-references: EMBL:X60592; NID:g29850; PIDN:CAA43045.1; PID:g29851
R:Braesch-Andersen, S.; Paulie, S.; Koho, H.; Niko, H.; Aspenstroem, P.; Perlmann, P.
J. Immunol. 142, 562-567, 1989
A:Title: Biochemical characteristics and partial amino acid sequence of the receptor-
A:Reference number: A60771; MUID:89093941
A:Accession: A60771
A:Molecule type: protein
A:Residues: 21-50 <BRA>
A:Experimental source: Burkitt lymphoma cell line Raji
C:Genetics:
A:Gene: GDB:CD40
A:Cross-references: GDB:215268; OMIM:109535
A:Map position: 20q12-20q13.2
C:Superfamily: CD27 antigen; NGF receptor repeat homology
C:Keywords: B-cell; glycoprotein; phosphoprotein; surface antigen; transmembrane prot
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-277/Product: B-cell activation protein CD40 #status experimental <MAT>
F:21-193/Domain: extracellular #status predicted <EXT>
F:194-215/Domain: transmembrane #status predicted <MM>
F:216-277/Domain: intracellular #status predicted <CYT>
F:153,180/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 12.9%; Score 211; DB 2; Length 277;
Best Local Similarity 30.5%; Pred. No. 7.6e-09;
Matches 46; Conservative 21; Mismatches 76; Indels 8; Gaps 3;
Qy 49 CAQCPGTFVQRCRRDSPTTCGCPCPRHRYTOFWNYLERC---RYCNVLGGEREEARAC 105
Db 38 CSLCQPGKLVSDCTFEFTETECPLCGESEFLDTWRETHCHQHKYCDPNLGRVQKGTGS 97
Qy 106 HATHNRACRCRTGFEFAH---GFCLHASCPPGAGVIAPGTPSQNTQCPCPPGTFSSASS 162

Db 98 ET--DTICTCEGWHCTSEACSVLHRSCTPGFVGKQIATGVSDTICEPCPGVGFSSNV 155

Qy 163 SSSEOCOPHRNCTALGLALNVPGSSSHDTLC 193

Db 156 SAFEKCHPWTSCTETKDLVVQQAQGNKTDVVC 186

RESULT 13

A46515

B cell-associated surface molecule CD40, short splice form - mouse

C:Species: Mus musculus (house mouse)

C:Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 26-May-2000

C:Accession: A46515

R:Grimaldi, J.C.; Torres, R.; Kozak, C.A.; Chang, R.; Clark, E.A.; Howard, M.; Cockayne, J. Immunol. 149, 3921-3926, 1992

A:Title: Genomic structure and chromosomal mapping of the murine CD40 gene.

A:Reference number: A46515; MUID:93094586

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: nucleic acid

A:Residues: 1-289 <GRI>

A:Cross-references: GB:M83312; NID:gl553058; PIDN:AAB08705.1; PID:gl553059; GB:M94126; N

A:Experimental source: BALB/c, liver

A:Note: sequence extracted from NCBI backbone (NCBIP:120357)

C:Comment: For an alternative splice form, see PIR:A46476

C:Superfamily: CD27 antigen; NGF receptor repeat homology

C:Keywords: alternative splicing

F:105-144/Domain: NGF receptor repeat homology <NGF>

Query Match 12.48; Score 203; DB 2; Length 289;

Best Local Similarity 31.18; Pred. No. 3.1e-08;

Matches 47; Conservative 19; Mismatches 77; Indels 8; Gaps 3;

Qy 49 CAQCPGTFVORPCRRDSTTCGPPRHYTQFWNYLERC---RYCNVLGGEREEERAC 105

Db 38 CDLCQPGSLTSHCTALEKTQCHPCDSGEFSAQWNREIRCHQHRCEPNOGLRVKKEGT- 96

Qy 106 HATHNRACRCRTGFFA---HAGFCLHASCPPGAGVIAPGTSONTOCQCPPTGFSASS 162

Db 97 -AESDTVCTCKEGQHCTSKDCEACQHTPCIPGFGVMEMATETTTDVTCHPCVPVGFSSNQS 155

Qy 163 SSSEOCOPHRNCTALGLALNVPGSSSHDTLC 193

Db 156 SLFEKCYPWTSCEDKNLEVLQKGTSTQTNVIC 186

RESULT 14

A46476

B cell-associated surface molecule CD40, long splice form - mouse

C:Species: Mus musculus (house mouse)

C:Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 26-May-2000

C:Accession: A46476

R:Torres, R.M.; Clark, E.A.

J. Immunol. 148, 620-626, 1992

A:Title: Differential increase of an alternatively polyadenylated mRNA species of murine

A:Reference number: A46476; MUID:92105763

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-305 <TOR>

A:Cross-references: GB:M83312; NID:gl553058

A:Note: sequence extracted from NCBI backbone (NCBIN:75206, NCBIP:75207)

A:Note: this translation is not annotated in Genbank entry MUSCD40A, release 113.0

C:Comment: For an alternative splice form, see PIR:A46515

C:Superfamily: CD27 antigen; NGF receptor repeat homology

C:Keywords: alternative splicing; Transmembrane protein

F:105-144/Domain: NGF receptor repeat homology <NGF>

Query Match 12.48; Score 203; DB 2; Length 305;

Best Local Similarity 31.18; Pred. No. 3.3e-08;

Matches 47; Conservative 19; Mismatches 77; Indels 8; Gaps 3;

Qy 49 CAQCPGTFVORPCRRDSTTCGPPRHYTQFWNYLERC---RYCNVLGGEREEERAC 105

Db 38 CDLCQPGSLTSHCTALEKTQCHPCDSGEFSAQWNREIRCHQHRCEPNOGLRVKKEGT- 96

Qy 106 HATHNRACRCRTGFFA---HAGFCLHASCPPGAGVIAPGTSONTOCQCPPTGFSASS 162

Db 97 -AESDTVCTCKEGQHCTSKDCEACQHTPCIPGFGVMEMATETTTDVTCHPCVPVGFSSNQS 155

Qy 163 SSSEOCOPHRNCTALGLALNVPGSSSHDTLC 193

Db 156 SLFEKCYPWTSCEDKNLEVLQKGTSTQTNVIC 186

RESULT 15

I48700

gene ox40 protein - mouse

N:Alternate names: Ox40 antigen

C:Species: Mus musculus (house mouse)

C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 11-Jan-2000

C:Accession: I48700; I48334; J34377

R:Calderhead, D.M.; Buhlmann, J.E.; van den Eertwegh, A.J.; Claassen, E.; Noelle, R.J. Immunol. 151, 5261-5271, 1993

A:Title: Cloning of mouse Ox40: a T cell activation marker that may mediate T-B cell

A:Reference number: I48700; MUID:94044750

A:Status: translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-272 <RES>

A:Cross-references: EMBL:Z21674; NID:g312827; PIDN:CAA79772.1; PID:g312828

R:Birkeland, M.L.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.; Barclay, A.N. Eur. J. Immunol. 25, 926-930, 1995

A:Title: Gene structure and chromosomal localization of the mouse homologue of rat Ox

A:Reference number: I48334; MUID:95255413

A:Accession: I48334

A:Status: translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-14, G', 16-272 <RE2>

A:Cross-references: EMBL:X85214; NID:g732818; PIDN:CAA59476.1; PID:g732819

C:Genetics: Ox40

A:Gene: Ox40

A:Introns: 45/1; 86/1; 122/1; 144/2; 210/1; 250/1

C:Superfamily: CD27 antigen; NGF receptor repeat homology

Query Match 12.18; Score 198.5; DB 2; Length 272;

Best Local Similarity 30.3%; Pred. No. 6.4e-08;

Matches 57; Conservative 17; Mismatches 71; Indels 43; Gaps 8;

Qy 15 VLALPALLPVPVAVRGVAETPTYPWRDAETGERLVCAQCPPTGVORPCRRDSTTCGCP 74

Db 11 LLLALLTLGLVTARLNCVKHTYP-----SGHK-CCRECQPGHGMVSRCDHTDRLTCLHPCE 64

Qy 75 PRHYTOFWNY--LERCYCNVLGGEREEERACHATHNRACRCRTGFFAHAGFCLHASC 132

Db 65 TGFYNEAVNYDTCKQCTQCNRHSG--SELKONCTPTQDTVCRCR----- 106

Qy 133 PPAGAVIAPGT-PSQNT-----QCQCPPTGTFSSSSSEOCOPHRNCTALGLALNVPG 185

Db 107 -----PGTQPRQDSGYKGLGVDCVCPGCHF--SPGNNAQACKPWTNCTLSGKQTRHEA 156

Qy 186 SSSSHDTLC 193

Db 157 SDSLDAVC 164

Search completed: January 30, 2001, 16:50:14

Job time: 704 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
OM protein - protein search, using sw model
Run on: January 30, 2001, 16:45:39 ; Search time 80.04 Seconds
(without alignments)
121.042 Million cell updates/sec
Title: US-09-518-931-2
Perfect score: 1634
Sequence: 1 MRALGPGLSLLCLLALPA.....RVARMPGLRSVRERFLPVH 300

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 88757 seqs, 32294092 residues
Total number of hits satisfying chosen parameters: 88757
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	Query Match	Length	ID	Description
1	353.5	21.6	461	1 TNR2_HUMAN	P20333 homo sapien
2	332.5	20.3	474	1 TNR2_MOUSE	P25119 mus musculus
3	315	19.3	435	1 TNR2_HUMAN	P36941 homo sapien
4	277	17.0	415	1 TNR2_MOUSE	P50284 mus musculus
5	262.5	16.1	349	1 VC22_VARV	P34015 variola vir
6	236.5	14.5	325	1 V722_SFVKA	P25943 shope fibro
7	226	13.8	277	1 OX40_HUMAN	P43489 homo sapien
8	215	13.2	326	1 V722_MYXVL	P29825 myxoma viru
9	214	13.1	271	1 OX40_RAT	P15725 rattus norv
10	211	12.9	277	1 CD40_HUMAN	P25942 homo sapien
11	203	12.4	289	1 CD40_MOUSE	P27512 mus musculus
12	202	12.4	269	1 CD40_BOVIN	O28203 bos taurus
13	198.5	12.1	272	1 OX40_MOUSE	P47741 mus musculus
14	186.5	11.4	595	1 CD30_HUMAN	P28908 homo sapien
15	185	11.3	256	1 41BB_MOUSE	P20334 mus musculus
16	176	10.8	416	1 NGFR_CHICK	P18519 gallus gall
17	175.5	10.7	427	1 NGFR_HUMAN	P08138 homo sapien
18	174	10.6	255	1 41BB_HUMAN	O07011 homo sapien
19	170	10.4	425	1 NGFR_RAT	P07174 rattus norv
20	163	10.0	332	1 FASA_PIG	O77736 sus scrofa
21	160.5	9.8	471	1 TNR1_BOVIN	O19131 bos taurus
22	155.5	9.5	260	1 CD27_HUMAN	P26842 homo sapien
23	155.5	9.5	327	1 FASA_MOUSE	P25446 mus musculus
24	155	9.5	323	1 FASA_BOVIN	P51867 bos taurus
25	148	9.1	250	1 CD27_MOUSE	P41272 mus musculus
26	145	8.9	335	1 FASA_HUMAN	P25445 homo sapien
27	144	8.8	324	1 FASA_RAT	O63199 rattus norv
28	137.5	8.4	454	1 TNR1_MOUSE	P25118 mus musculus
29	128.5	7.9	417	1 WSL1_HUMAN	O93038 h wsl-1 pro
30	128.5	7.9	3635	1 LMA5_MOUSE	O61001 mus musculus
31	127	7.8	1192	1 LMG2_MOUSE	O61092 mus musculus
32	125.5	7.7	461	1 TNR1_RAT	P22934 rattus norv
33	125.5	7.7	1255	1 ERB2_HUMAN	P04626 homo sapien

34	124.5	7.6	1713	1 LMA3_HUMAN	Q16787 homo sapien
35	124.5	7.6	3106	1 LMA2_MOUSE	O60675 mus musculus
36	123.5	7.6	455	1 TNR1_HUMAN	P19438 homo sapien
37	123.5	7.6	2813	1 VWF_CANFA	Q28295 canis famil
38	123	7.5	2569	1 LMA3_MOUSE	Q61789 mus musculus
39	120	7.3	1609	1 LMGI_HUMAN	P11047 homo sapien
40	120	7.3	2318	1 TNR3_MOUSE	Q61982 mus musculus
41	119.5	7.3	461	1 TNR1_PIG	P50555 sus scrofa
42	119.5	7.3	3084	1 LMA1_MOUSE	P19137 mus musculus
43	118.5	7.3	1680	1 FUR2_DROME	P30432 drosophila
44	117.5	7.2	1964	1 NTC4_MOUSE	P31695 mus musculus
45	116.5	7.1	3110	1 LMA2_HUMAN	P24043 homo sapien

ALIGNMENTS

RESULT 1
TNR2_HUMAN
ID TNR2_HUMAN STANDARD; PRT; 461 AA.
AC P20333;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR 2 PRECURSOR (TUMOR NECROSIS FACTOR
DE BINDING PROTEIN 2) (TBPII) (P80) (TNF-R2) (P75) (CD120B) (ETANERCEPT).
GN TNFRSF1B OR TNFR2 OR TNFR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE-90260639; PubMed-2160731;
RA Smith C.A., Davis T., Anderson D., Solam L., Beckmann M.P., Jerzy R.,
RA Dower S.K., Cosman D., Goodwin R.G.;
RT "A receptor for tumor necrosis factor defines an unusual family of
RT cellular and viral proteins.";
RL Science 248:1019-1023(1990).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE-91045991; PubMed-2172983;
RA Kohno T., Brewer M.T., Baker S.L., Schwartz P.E., King M.W.,
RA Hale K.K., Squires C.H., Thompson R.C., Vannice J.L.;
RT "A second tumor necrosis factor receptor gene product can shed a
RT naturally occurring tumor necrosis factor inhibitor.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:8331-8335(1990).
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE-96299745; PubMed-86611109;
RA Beltlinger C.P., White P.S., Maris J.M., Sulman E.P., Jensen S.J.,
RA Lepaslier D., Stallard B.J., Goeddel D.V., Desauvage F.J.,
RA Brodeur G.M.;
RT "Physical mapping and genomic structure of the human TNFR2 gene.";
RL Genomics 35:94-100(1996).
[4]
RP SEQUENCE OF 116-461 FROM N.A. AND PARTIAL SEQUENCE.
RX MEDLINE-90349572; PubMed-2166946;
RA Heller R.A., Song K., Onasch M.A., Fischer W.H., Chang D.,
RA Ringold G.M.;
RT "Complementary DNA cloning of a receptor for tumor necrosis factor
RT and demonstration of a shed form of the receptor.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:6151-6155(1990).
[5]
RP SEQUENCE OF 27-31.
RX MEDLINE-90110215; PubMed-2153136;
RA Engelmann H., Novick D., Wallach D.;
RT "Two tumor necrosis factor-binding proteins purified from human
RT urine. Evidence for immunological cross-reactivity with cell surface
RT tumor necrosis factor receptors.";
RL J. Biol. Chem. 265:1531-1536(1990).
[6]
RP SEQUENCE OF 22-40; 65-69; 136-141; 300-306 AND 346-362.
RX MEDLINE-91056048; PubMed-2173696;

Loetscher H., Schlaeger E.J., Lahm H.-W., Pan Y.-C.E., Lesslauer W., Brockhaus M.;
"Purification and partial amino acid sequence analysis of two distinct tumor necrosis factor receptors from HL60 cells."; J. Biol. Chem. 265:20131-20138(1990);

[7]
CHARACTERIZATION.
MEDLINE=93016040; PubMed=1328224;
Pennica D., Lam V.T., Mize N.K., Weber R.F., Fendly B.M., Lipari M.T., Goeddel D.V.;
"Biochemical properties of the 75-kDa tumor necrosis factor receptor. Characterization of ligand binding, internalization, and receptor phosphorylation."; J. Biol. Chem. 267:21172-21178(1992).

[8]
X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 419-428 IN TRAF2 COMPLEX.
MEDLINE=99221490; PubMed=10206649;
Park Y.C., Burkitt V., Villa A.R., Tong L., Wu H.;
"Structural basis for self-association and receptor recognition of human TRAF2."; Nature 398:533-538(1999).
-!- FUNCTION: RECEPTOR FOR TNF-ALPHA. HIGH AFFINITY FOR TNA-ALPHA AND APPROXIMATELY 5-FOLD LOWER AFFINITY FOR TNF-BETA.
-!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
-!- PTM: PHOSPHORYLATED; MAINLY ON SERINE RESIDUES WITH A VERY LOW LEVEL ON THREONINE RESIDUES.
-!- PHARMACEUTICAL: AVAILABLE UNDER THE NAME ENBREL (IMMUNEX AND WYETH-AYEAST). USED TO TREAT MODERATE TO SEVERE RHEUMATOID ARTHRITIS (RA). ENBREL CONSIST OF THE EXTRACELLULAR LIGAND-BINDING PORTION OF TNFR2 LINKED TO AN IMMUGLOBULIN FC CHAIN. IT BINDS TO TNF-ALPHA AND BLOCKS ITS INTERACTIONS WITH RECEPTORS.
-!- SIMILARITY: CONTAINS A LA-NGR/TNFR-TYPE CYSTEINE-RICH REGION.
-!- DATABASE: NAME=PROW; NOTE=CD guide CD120b entry;
WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd120b.htm".
-!- DATABASE: NAME=Enbrel; NOTE=Clinical information on Enbrel;
WWW="http://www.enbrelinfo.com/";

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M32315; AAC59929.1; -.
EMBL; M35857; AAA63262.1; -.
EMBL; U52165; AAC50622.1; -.
EMBL; U52156; AAC50622.1; JOINED.
EMBL; U52157; AAC50622.1; JOINED.
EMBL; U52158; AAC50622.1; JOINED.
EMBL; U52159; AAC50622.1; JOINED.
EMBL; U52160; AAC50622.1; JOINED.
EMBL; U52161; AAC50622.1; JOINED.
EMBL; U52162; AAC50622.1; JOINED.
EMBL; U52163; AAC50622.1; JOINED.
EMBL; U52164; AAC50622.1; JOINED.
EMBL; M55994; AAA36755.1; -.
PIR; A35356; A35356.
PIR; A36007; A36007.
PIR; A36475; A36475.
PIR; B35010; B35010.
PIR; A23666; A23666.
PDB; 1CA9; 12-APR-99.
MIM; 191191; -.
INTERPRO; IPRO01368; -.
PFAM; PF00020; TNFR_C6; 4.
PROSITE; PS00652; TNFR_NGFR_1; 2.
PROSITE; PS00050; TNFR_NGFR_2; 4.
Receptor; Transmembrane; Glycoprotein; Repeat; Signal;
Phosphorylation; Pharmaceutical; 3D-structure.
SIGNAL 1
22

FT DOMAIN 23 257 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 258 287 POTENTIAL.
FT DOMAIN 288 461 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 39 201 4 X TNFR-CYS.
FT REPEAT 39 76 TNFR-CYS 1.
FT REPEAT 77 118 TNFR-CYS 2.
FT REPEAT 119 162 TNFR-CYS 3.
FT REPEAT 163 201 TNFR-CYS 4.
FT DISULFID 40 53 BY SIMILARITY.
FT DISULFID 54 67 BY SIMILARITY.
FT DISULFID 57 75 BY SIMILARITY.
FT DISULFID 78 93 BY SIMILARITY.
FT DISULFID 96 110 BY SIMILARITY.
FT DISULFID 100 118 BY SIMILARITY.
FT DISULFID 120 126 BY SIMILARITY.
FT DISULFID 134 143 BY SIMILARITY.
FT DISULFID 137 161 BY SIMILARITY.
FT CARBOHYD 171 171 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 193 193 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 141 141 R -> P (IN REF. 4).
FT CONFLICT 196 196 R -> M (IN REF. 1 AND 3).
FT CONFLICT 363 363 A -> T (IN REF. 4).
SQ SEQUENCE 461 AA; 48316 MW; 603B580ECD67636F CRC64;

Query Match 21.6%; Score 353.5; DB 1; Length 461;
Best Local Similarity 29.8%; Pred. No. 3.8e-20;
Matches 96; Conservative 43; Mismatches 122; Indels 61; Gaps 12.

QY 8 GLSLCLLVALLPALLPVAVRGVAETPTYPWRDAETGE-----RVLCACQPPG 55
||| |||| | | | | | | | | | | | | | | | | : | : | : |
Db 13 GLELWAAHALPA-----QVATTPYAP----EPGSTCLRLEYDQTAEOMCKSCSPG 60

QY 56 TFVQRPCRDRSDPTTCGPPRPRTYTQFWNVYLERCYNVLCGEREEARACHATHNRACRC 115
||| |||| | | | | | | | | | | | | | | | | : | : | | |
Db 61 QHAKVFCTKTSVTDCSEDSYTTQLWNVPCLSCGSRCSDDOVETOACTREDNICTC 120

QY 116 RTGFFAHAG-----FCLEHASCPGGAGVIAPGTQSONTCQCPCPGTSSASSSEQCQ 169
||| |||| | | | | | | | | | | | | | | | | : | : | : | : | : | : |
Db 121 RPQWYCALSQKECRCLCAPLRKCRPGFGVARPGTETSDVVCKPCAPGTFSNTSSDICR 180

QY 170 PHRNCTALGIALNVPGSSSHDTLTCTCTGFPLSTRVPRGAECERAVIDFVAFODISIKRL 229
||| |||| | | | | | | | | | | | | | | | | : | : | : | : | : | : |
Db 181 PFOICNVVA---IPGNASRDVCTSTS---PTRSMAPGAVHLDPQV-----STRSOHT 227

QY 230 QRLLQALEAPE-----GWGPTPRA---GRAALQLKLRRRLTELGAQGADGALLVRLLQAL 280
||| |||| | | | | | | | | | | | | | | | | : | : | : | : | : | : |
Db 228 QPTPEPSTAPSTFLIPMGSPSPAEGSTGDFALPVGLIVGVITAL-----GLLIIGVNCV 282

QY 281 ---RVARMP--GLERSVNERFLP 298
||| |||| | | | | | | | | | | | | | | | | : | : | : | : | : | : |
Db 283 IMTVKKKKLCLQREAKVPHLP 304

RESULT 2
TNFR2_MOUSE STANDARD; PRT; 474 AA.
ID TNFR2_MOUSE
AC P25119; P97893;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR 2 PRECURSOR (TNF-R2) (P75).
GN TNFRSF1B OR TNFR2 OR TNFR-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
[1]
SEQUENCE FROM N.A.
RP MEDLINE=91187885; PubMed=1849278;
RX Lewis M., Tartaglia L.A., Lee A., Bennett G.L., Rice G.C.,
RA Wong G.H., Chen E.Y., Goeddel D.V.;
RA "Cloning and expression of cDNAs for two distinct murine tumor
FT SIGNAL 1
22

RT necrosis factor receptors demonstrate one receptor is species
 RL specific.";
 RN Proc. Natl. Acad. Sci. U.S.A. 88:2830-2834(1991).
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91246168; PubMed=1645445;
 RA Goodwin R.G., Jenkins D., Jerzy R., Davis T., Brannan C.I.,
 RA Copeland N.G., Jenkins N.A., Smith C.A.;
 RT "Molecular cloning and expression of the type 1 and type 2 murine
 RT receptors for tumor necrosis factor.";
 RL Mol. Cell. Biol. 11:3020-3026(1991).
 RN [3]
 RN SEQUENCE OF 1-26 FROM N.A.
 RC STRAIN=NOD;
 RA Jacob C.O., Liu J.;
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RN SEQUENCE OF 1-22 FROM N.A.
 RC TISSUE=LIVER;
 RA Kissnerghis M., Fellows R., Feldmann M., Chernajovsky Y.;
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: RECEPTOR FOR TNF-ALPHA.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M60469; AAA39752.1; -;
 DR EMBL; M59378; AAA40463.1; -;
 DR EMBL; U39488; AAA85021.1; -;
 DR EMBL; X87128; CAA60618.1; -;
 DR PIR; B38634; B38634;
 DR HSP; P19438; INCF.
 DR MGD; MGI:1314883; TNFRSF1B.
 DR INTERPRO; IPR001368; -;
 DR PFAM; PF00020; TNFR_C6; 4.
 DR PROSITE; PS00652; TNFR_NGFR_1; 2.
 DR PROSITE; PS00500; TNFR_NGFR_2; 3.
 DR Receptor; Transmembrane; Glycoprotein; Repeat; Signal.
 FT SIGNAL
 FT CHAIN 1 22
 FT DOMAIN 23 474 TUMOR NECROSIS FACTOR RECEPTOR 2.
 FT TRANSMEM 23 258 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 259 288 POTENTIAL.
 FT DOMAIN 289 474 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 39 203 4 X TNFR-CYS.
 FT REPEAT 39 77 TNFR-CYS 1.
 FT REPEAT 78 119 TNFR-CYS 2.
 FT REPEAT 120 164 TNFR-CYS 3.
 FT REPEAT 165 203 TNFR-CYS 4.
 FT DISULFID 40 54 BY SIMILARITY.
 FT DISULFID 55 68 BY SIMILARITY.
 FT DISULFID 58 76 BY SIMILARITY.
 FT DISULFID 79 94 BY SIMILARITY.
 FT DISULFID 97 111 BY SIMILARITY.
 FT DISULFID 101 119 BY SIMILARITY.
 FT DISULFID 121 127 BY SIMILARITY.
 FT DISULFID 136 145 BY SIMILARITY.
 FT DISULFID 139 163 BY SIMILARITY.
 FT DISULFID 166 181 BY SIMILARITY.
 FT CARBOHYD 69 69 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 195 195 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 474 AA; 50319 MW; 462EAE398C4D6563 CRC64;

Query Match 20.3%; Score 332.5; DB 1; Length 474;
 Best Local Similarity 29.7%; Pred. No. 1.6e-18;
 Matches 81; Conservative 44; Mismatches 109; Indels 39; Gaps 9;

Qy 46 RLVAQCPTGTFVQPCRDSPPTTCGPPRHYTOFWNLRCRYCNVLGCEEEARAC 105
 Db 52 QMCCAKPPGQYVHFCKNTSDTVACDCEASMTYVWNOFTCLSCSSCTTDQVEIRAC 111
 Qy 106 HATHNRACRRTGFF-----AHAGF---CLEHASCPPGAGVIAPGTPSQNTQCPCPPGTF 158
 Db 112 TKQNRVCACEAGRYCALKTHSGSCRCQMLSKCGPGFVASSRAPNGNVLCKACAPGTF 171
 Qy 159 SASSSSSQCPHNRCTALGLALNVGSSSHDTLCT-----SCTGPPLSTRVPGABECERA 214
 Db 172 SDTTSSTDVCRPHRTCSILA-----IPGNASTDAVCAPESTLSAIPRTLYVSQPEPTRSQ 227
 Qy 215 VDFVAFODISIKRLQRLQALAEPEGPTP-----RAGRAALQKLRRRLTELLGAQD 269
 Db 228 PLD-----QEPGSPQPSILTSU-----GSTPIISQTKGGISLPIGLIVGVTSL----- 272
 Qy 270 GALLVRLQLAL-----RVARMPGLSRRERFLP 298
 Db 273 GLLMLGLVNCILVQRKKKPSCLQRDAKVPHVP 305
 RESULT 3
 TNRC_HUMAN
 ID TNRC_HUMAN STANDARD; PRT; 435 AA.
 AC P36941;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE LYMPHOTOXIN-BETA RECEPTOR PRECURSOR (TUMOR NECROSIS FACTOR RECEPTOR
 DE 2 RELATED PROTEIN) (TUMOR NECROSIS FACTOR C RECEPTOR).
 GN LTBR OR TNFCR OR TNFRSF3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE=93252381; PubMed=8486360;
 RA Baens M., Chaffanet M., Cassiman J.J., den Berghe H., Marynen P.;
 RT "Construction and evaluation of a hncDNA library of human l2p
 RT transcribed sequences derived from a somatic cell hybrid.";
 RL Genomics 16:214-218(1993).
 RN [2]
 RP FUNCTION.
 RX MEDLINE=94225209; PubMed=8171323;
 RA Crowe P.D., van Arsdale T.L., Walter B.N., Ware C.F., Hession C.,
 RA Ehrenfels B., Browning J.L., Din W.S., Goodwin R.G., Smith C.A.;
 RL "A lymphotoxin-beta-specific receptor.";
 RL Science 264:707-710(1994)
 CC -1- FUNCTION: RECEPTOR FOR THE LYMPHOTOXIN-BETA. POSSIBLE FUNCTION IN
 CC IMMUNE DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L04270; AAA36757.1; -;
 DR HSSP; P25942; ICDF.
 DR MIM; 600979; -;
 DR INTERPRO; IPR001368; -;
 DR PFAM; PF00020; TNFR_C6; 4.
 DR PROSITE; PS00652; TNFR_NGFR_1; 2.
 DR PROSITE; PS00500; TNFR_NGFR_2; 3.
 FT SIGNAL 1 30
 FT CHAIN 31 435 LYMPHOTOXIN-BETA RECEPTOR.

Wed Jan 31 13:43:51 2001

```

FT DOMAIN 31 227 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 228 248 POTENTIAL.
FT DOMAIN 249 435 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 42 211 4 X TNFR-CYS.
FT REPEAT 42 81 TNFR-CYS 1.
FT REPEAT 82 124 TNFR-CYS 2.
FT REPEAT 125 168 TNFR-CYS 3.
FT REPEAT 169 211 TNFR-CYS 4.
FT DISULFID 43 58 BY SIMILARITY.
FT DISULFID 59 72 BY SIMILARITY.
FT DISULFID 62 80 BY SIMILARITY.
FT DISULFID 83 98 BY SIMILARITY.
FT DISULFID 101 116 BY SIMILARITY.
FT DISULFID 104 124 BY SIMILARITY.
FT DISULFID 126 132 BY SIMILARITY.
FT DISULFID 139 148 BY SIMILARITY.
FT DISULFID 142 167 BY SIMILARITY.
FT DISULFID 170 185 BY SIMILARITY.
FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 435 AA; 46709 MW; 624626B5022F656F CRC64;

Query Match 19.38; Score 315; DB 1; Length 435;
Best Local Similarity 31.88; Pred. No. 3.1e-17;
Matches 89; Conservative 29; Mismatches 120; Indels 42; Gaps 12;

Qy 3 ALEPGSLLLCLVALLPALLPVPVAVGVAETPTT-----PWRDA-----ETGERLVCAQC 52
Dy 6 ATSPGLAWGLVLGLFLGLAASQPAV---PPYASENQTCDQOEKEYEYEPQHRIICSRC 62

Qy 53 PPGTIVQPCRRDPTTCGCPGPPRHYYTFWNLYL---EPCRYCNVLCGEREEERACHATH 109
Dy 63 PPGTIVSAKCSKIRDTVCATCAENSYNHWNLYTICQLCRPCDPVNG--LEEIAPCTSKR 120

Qy 110 NRACRCRTGFFAHAGFLE--H---ASCPPGA-GVIAPGTPSPQTCQPCPGPTFSASS 162
Dy 121 KTQCRQCPGMEC-AAWALECTHCELLSDCPGTEALKEDEVGKGNHNCVPCKAGHFONT 179

Qy 163 SSSEQCPHRCNTALGLALNVPGSSSHDTLTCTGTPPLSTRVPGAECECAVIDFVAQ 222
Dy 180 SPARCPQHTRCENGGLVEAAPTGAQSDTCKNPLE-PLPPMSGTMLMLAVLLPLAFFL 238

Qy 223 DIS-----IKRLQLLOALEAPGCGPTPRAG 249
Dy 239 LLATVPSCITKWSHPSLCRLKIGSLK--RRPQEGENPVAG 276

RESULT 4
TNRC_MOUSE STANDARD; PRT; 415 AA.
AC P50284;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE LYMPHOTOXIN-BETA RECEPTOR PRECURSOR.
GN LTBR OR TNFR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=CVB; TISSUE=LUNG;
RX MEDLINE=96072804; PubMed=7594541;
RA Force W.R., Walter B.N., Hession C., Tizard R., Kozak C.A.,
RA Browning J.L., Ware C.F.;
RT "Mouse lymphotoxin-beta receptor. Molecular genetics, ligand binding,
and expression."
RL J. Immunol. 155:5280-5288(1995).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=96163885; PubMed=8586432;
RA Nakamura T., Tashiro K., Nazarea M., Nakano T., Sasayama S.,

```

```

RA Honjo T.;
RT "The murine lymphotoxin-beta receptor cDNA: isolation by the signal
sequence trap and chromosomal mapping.";
RL Genomics 30:312-319(1995). THE LYMPHOTOXIN-BETA. POSSIBLE FUNCTION IN
CC -I- FUNCTION: RECEPTOR FOR THE LYMPHOTOXIN-BETA. IMMUNE DEVELOPMENT.
CC -I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -I- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U29173; AAA68964.1; -
CC EMBL; L38423; AAA00846.1; -
CC DR EMBL; U30798; AAA81334.1; -
CC DR HSP; P25942; ICDP.
CC DR MGD; MGI:104875; LTBR.
CC INTERPRO; IPR001368; -
CC PFAM; PF00020; TNFR_C6; 3.
CC PROSITE; PS00652; TNFR_NGFR_1; 2.
CC DR PROSITE; PS00500; TNFR_NGFR_2; 3.
CC KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 30 POTENTIAL.
FT CHAIN 31 415 LYMPHOTOXIN-BETA RECEPTOR.
FT DOMAIN 31 223 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 224 244 POTENTIAL.
FT DOMAIN 245 415 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 42 213 4 X TNFR-CYS.
FT REPEAT 42 81 TNFR-CYS 1.
FT REPEAT 82 124 TNFR-CYS 2.
FT REPEAT 125 170 TNFR-CYS 3.
FT REPEAT 171 213 TNFR-CYS 4.
FT DISULFID 43 58 BY SIMILARITY.
FT DISULFID 59 72 BY SIMILARITY.
FT DISULFID 62 80 BY SIMILARITY.
FT DISULFID 83 98 BY SIMILARITY.
FT DISULFID 101 116 BY SIMILARITY.
FT DISULFID 104 124 BY SIMILARITY.
FT DISULFID 126 132 BY SIMILARITY.
FT DISULFID 139 150 BY SIMILARITY.
FT DISULFID 142 169 BY SIMILARITY.
FT DISULFID 172 187 BY SIMILARITY.
FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 415 AA; 44956 MW; 29B326A566AEF661 CRC64;

Query Match 17.08; Score 277; DB 1; Length 415;
Best Local Similarity 31.2%; Pred. No. 2.4e-14;
Matches 88; Conservative 29; Mismatches 105; Indels 60; Gaps 14;

Qy 6 GP---GLSLCLVLLPALLPVPVAVGVAETPTTPWRDAETGERL---VCAQCPPGPTFVQ 59
Dy 15 GPLLLGLSLGL-LVASQQLVPPYRI-----ENQTCWDQDKEYEYEPMDVCCSRCPGEFV 69

Qy 60 RPCRDRSPTTCGCPGPPRHYYTFWNLYL---EPCRYCNVLCGEREEERACHATHINRACR 116
Dy 70 AVCSRSQDVTCKTCPHNSYNEHWNHLSLTCQLCRPCDIVLG--FEVAPCTSDRKAECRC 127

Qy 117 TGFFAHAGFCL---EHASCPGAGVIA-PCTPSQNT-----OCPCPPGPTFSASS 163
Dy 128 PGM-----SCVYLDNECVHCEERLVLCQPGTEAEVDEIMDTDVNCVCKPGHFWTSS 182

Qy 164 SSEQCPHRCNTALGLALNVPGSSSHDTLTCTGTPPLSTRVPGAECECAVIDFVAF-- 221
Dy 183 PRACQPHTRCEIQGLVEAAGTYSYSDTICK-----NPPEGAMLLAILLSLVFL 235

Qy 222 -----QDTSIKRLQLLOALEAPGCGW---PTPRA 248

```

RESULT 6

125 FCLHASCPGAGVIAPGTSPNTQCCPPGTTSASSSSSEQCQHNCALNALNVP 184
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
122 ICAPQTKPCAGYGV-SGHTRAGDTLCEKCPHTYSDSLSPTERCGTSPNYISVGFNL--- 177

185 GSSSHDITLTCTSG 198
::| ||: |
178 -YPVNETSCTTTAG 190

RESULT 7
OX40_HUMAN OX40_HUMAN STANDARD; PRT; 277 AA.
AAC P43489; Q13663;
01-NOV-1995 (rel. 32, Created)
01-NOV-1995 (rel. 32, Last sequence update)
01-OCT-2000 (rel. 40, Last annotation update)
OX40L RECEPTOR PRECURSOR (ACT35 ANTIGEN) (TAX-TRANSCRIPTIIONALLY
DE ACTIVATED GLYCOPROTEIN 1 RECEPTOR) (CD134 ANTIGEN).
GN TNFRSF4 OR TXGP1L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RZ SEQUENCE FROM N.A.
RX MEDLINE=94170844; PubMed=7510240;
RA Latza U., Duerkop H., Schnittger S., Ringeling J., Eitelbach F.,
Hummel M., Fonatsch C., Stein H.;
"The human OX40 homolog: cDNA structure, expression and chromosomal
assignment of the ACT35 antigen.";
RT Eur. J. Immunol. 24:677-683(1994).
RL [2]
RZ SEQUENCE FROM N.A.
RX MEDLINE=95219871; PubMed=7704935;
RA Baum P.R., Gayle R.B. III, Ramsdell F., Srinivasan S., Sorensen R.A.,
Watson M.L., Seldin M.F., Clifford K.N., Grabstein K., Alderson M.R.;
"Identification of OX40 ligand and preliminary characterization of
its activities on OX40 receptor.";
RL Circ. Shock 44:30-34(1994).
RZ FUNCTION: RECEPTOR FOR THE OX40L/GP34 CYTOKINE.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: CONTAINS A LA-NGRF/TNFR-TYPE CYSTEINE-RICH REGION.
CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD134 entry;
WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd134.htm".

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@lsb-sib.ch).

EMBL; X75962; CAAS3576.1; -
DR EMBL; S76792; AAB33944.1; ALT_INIT.
DR HSSP; P25942; LCDF.
DR MIM; 600315; -
DR INTERPRO; IPR001368; -
DR PFAM; PF00020; TNFR_C6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; 3.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
KW Receptor; T-cell; Antigen; Glycoprotein; Transmembrane; Repeat;
Signal.
FT SIGNAL 1 28
FT CHAIN 29 277
FT DOMAIN 29 214
FT TRANSMEM 215 235
FT DOMAIN 216 277
FT DOMAIN 30 169
FT REPEAT 30 65
FT REPEAT 66 107
FT REPEAT 108 126
FT REPEAT 127 167
FT POTENTIAL.
FT OX40L RECEPTOR..
FT EXTRACELLULAR {POTENTIAL}..
FT POTENTIAL.
FT CYTOPLASMIC {POTENTIAL}..
FT 4 X TNFR-CYS.
FT TNFR-CYS 1.
FT TNFR-CYS 2.
FT TNFR-CYS 3 {INCOMPLETE}.
FT TNFR-CYS 4.

```
FT CHAIN 17 326 TUMOR NECROSIS FACTOR SOLUBLE RECEPTOR.
FT DOMAIN 27 186 4 X TNFR-CYS.
FT REPEAT 27 62 TNFR-CYS 1.
FT REPEAT 63 104 TNFR-CYS 2.
FT REPEAT 105 147 TNFR-CYS 3.
FT REPEAT 148 186 TNFR-CYS 4.
FT CARBOHYD 66 66 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 326 AA; 35208 MW; ABBF027E947292FF CRC64;

Query Match 13.2%; Score 215; DB 1; Length 326;
Best Local Similarity 29.3%; Pred. No. 1e-09;
Matches 58; Conservative 22; Mismatches 96; Indels 22; Gaps 5;

Qy 12 LCLVLALPALL-----PVPARGVAETPTYPWRDAETGERLVCAQCPGPTFVQRCRRDS 66
Db 4 LTLLEYVACVYGGAGYGDAGKGRNDY-----EKDGLCTSCPPGVSASRLCGPGS 57
Qy 67 PTCGCPPHRYTQFWNYLERCVNVLGCREREEARACHATHNRACRGTGFFA----- 121
Db 58 DTVGSPCKNETFTASTNAPACVSCGRCTGHLSESQCDKTRDRVDCSAGNYCLLRGQ 117
Qy 122 -HAGFCLEHASCPPGAGVIAPGT-PSQNT-----QCQCPGPTGTFSSSSSEQCQPHRNCNTALGLA 180
Db 118 EGCRCAPKTKPCAGYGV-SGHTRTGVDLCTKCPRYTSDAVSSTETCTSSFNYSVEFN 176
Qy 181 LNVPGSSSHDLCTSGT 198
Db 177 L-----YPVNDTSCTTAG 190

RESULT 9
OX40_RAT
ID OX40_RAT STANDARD; PRT; 271 AA.
AC P15725;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE OX40L RECEPTOR PRECURSOR (OX40 ANTIGEN) (MRC OX40).
GN TNFRSF4 OR TXGPIL OR OX40.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-T-CELL;
RX MEDLINE=90214614; PubMed=2157591;
RA Mallett S., Fossum S., Barclay A.N.;
RT "Characterization of the MRC OX40 antigen of activated CD4 positive T
RT lymphocytes - a molecule related to nerve growth factor receptor.";
RL EMBO J. 9:1063-1068(1990).
CC -!- FUNCTION: RECEPTOR FOR THE OX40L/GP34 CYTOKINE.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: ACTIVATED T-CELLS.
CC -!- SIMILARITY: CONTAINS A LA-NGR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X17037; CAA34897.1; -.
DR PIR; S08036; S08036.
DR PIR; S12783; S12783.
DR HSSP; P25942; IGF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_c6; 3.
```

```
DR PROSITE; PS00652; TNFR_NGFR_1; 3.
DR PROSITE; PS00500; TNFR_NGFR_2; 2.
KW Receptor; T-cell; Antigen; Glycoprotein; Transmembrane; Repeat;
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 271 OX40L RECEPTOR.
FT DOMAIN 20 210 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 211 235 POTENTIAL.
FT DOMAIN 236 271 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 25 164 4 X TNFR-CYS.
FT REPEAT 25 60 TNFR-CYS 1.
FT REPEAT 61 102 TNFR-CYS 2.
FT REPEAT 103 123 TNFR-CYS 3 (INCOMPLETE).
FT REPEAT 124 164 TNFR-CYS 4.
FT CARBOHYD 143 143 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 271 AA; 29895 MW; C0645136B16E821 CRC64;

Query Match 13.1%; Score 214; DB 1; Length 271;
Best Local Similarity 30.1%; Pred. No. 1e-09;
Matches 58; Conservative 23; Mismatches 64; Indels 48; Gaps 9;

Qy 10 SLLCLVLALPALLPVPARGVAETPTYPWRDAETGERLVCAQCPGPTFVQRCRRDSPTT 69
Db 10 AFLLLGLSLGVTVKLNCVK-----DTYP-----SGHK-CCRECQPGHGMVSRCDHTRDTV 58
Qy 70 CGCPCPHRYTQFWNY--LERCVNVLGCREREEARACHATHNRACRGTGFFAHAGFCL 127
Db 59 CHCPEPGYNAVNVDTCQCTQCNRSG--SELKQNTPTEDTVCCQCR----- 105
Qy 128 EHACPPGAGVIAPGT-PSQNT-----QCQCPGPTGTFSSSSSEQCQPHRNCNTALGLA 180
Db 106 -----PGTQPRQDSSHLKLGVCVPCPPGHE--SPGSNQACKPWNTNCTLSGKQ 150
Qy 181 LNVPGSSSHDLCT 193
Db 151 IRHPASNSLDTVC 163

RESULT 10
CD40_HUMAN
ID CD40_HUMAN STANDARD; PRT; 277 AA.
AC P25942;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CD40L RECEPTOR PRECURSOR (B-CELL SURFACE ANTIGEN CD40) (BP50) (CDW40)
DE (TUMOR NECROSIS FACTOR RECEPTOR 5).
GN TNFRSF5 OR CD40.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89355608; PubMed=2475341;
RA Stamenkovic I., Clark E.A., Seed B.;
RT "A B-lymphocyte activation molecule related to the nerve growth
RT factor receptor and induced by cytokines in carcinomas.";
RL EMBO J. 8:1403-1410(1989).
RN [2]
RN 3D-STRUCTURE MODELING OF 24-144.
RX MEDLINE=97189482; PubMed=9037712;
RA Bajorath J., Aruffo A.;
RT "Construction and analysis of a detailed three-dimensional model of
RT the ligand binding domain of the human B cell receptor CD40.";
RL Proteins 27:59-70(1997).
RN [3]
RN 3D-STRUCTURE MODELING OF 26-186 IN COMPLEX WITH CD40L.
RX MEDLINE=98266353; PubMed=9605317;
RA Singh J., Garber E., van Vlijmen H., Karpusas M., Hsu Y.-M.,
RA Zheng Z., Naismith J.H., Thomas D.;
RT "The role of polar interactions in the molecular recognition of CD40L
RT with its receptor CD40.";
```


Query Match 12.9%; Score 211; DB 1; Length 277;
Best Local Similarity 30.5%; Pred. No. 1.7e-09;
21. Mismatches 76; Indels 8; Gaps 3;

RESULT	11
CD40_MOUSE	
ID	CD40_MOUSE
AC	P27512
DT	01-AUG-1992
DT	01-OCT-1996
DT	Rel. 34, Last sequence update)
DT	15-JUN-1999
DT	Rel. 38, Last annotation update)
STANDARD:	
PRT:	289 AA.

Query Match 12.4%; Score 203; DB 1; Length 289;
Best Local Similarity 31.1%; Pred. No. 7.4e-09;

Qy 49 CAQCPPTGVOPCRDRSPITTCGCPRHRYTQFWNYLERC---RYCNVLGGEREEEARAC 105
- - - - -
Db 38 CDLCQPGSRLTSHCTALEKTQCHPCDSGEFSAQWNREIRCHQHRRCEPNGLRVKKEGT- 96
- - - - -
Qy 106 HATHNRACRRTGTFGA---HAGTCLHBASCPGAGVIAPGTSPQNTQCPQCPPTFSASS 162


```
Db 97 -AESDTVCTCKEQHCTSKDCACQAHTPCIPGFGVMEMATTETDTVCHPCPVGFESNQS 155
QY 163 SSSSECOQPHNCTALGLALNVPSSSHDFLC 193
Db 156 SLFEKCPWTSCEKDNLEVLQKGTSTQTNVIC 186

RESULT 12
CD40_BOVIN STANDARD; PRT; 269 AA.
ID CD40_BOVIN
AC Q28203;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CD40L RECEPTOR PRECURSOR (B-CELL SURFACE ANTIGEN CD40) (FRAGMENT).
GN TNFRSF5 OR CD40.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-97281252; PubMed-9135560;
RA Hirano A., Brown W.C., Estes D.M.;
RT "Cloning, expression and biological function of the bovine CD40
RT homologue: role in B-lymphocyte growth and differentiation in
RT cattle";
RL Immunology 90:294-300(1997).
RC -!- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD40L.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -!-
DR EMBL; U57745; AAC48710.1; -.
DR HSSP; P25942; ICDF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 1.
DR PROSITE; PS50050; TNFR_NGFR_2; 1.
KW Receptor; B-cell; Glycoprotein; Transmembrane; Repeat; Signal.
FT SIGNAL 1 19
FT CHAIN 20 >269 CD40L RECEPTOR.
FT DOMAIN 20 193 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 194 215
FT DOMAIN 216 >269
FT CYTOPLASMIC (POTENTIAL).
FT DOMAIN 25 187
FT TNFR-CYS 1.
FT REPEAT 25 60 TNFR-CYS 2.
FT REPEAT 61 103 TNFR-CYS 3.
FT REPEAT 104 144 TNFR-CYS 4.
FT REPEAT 145 187
FT CARBOHYD 153 153 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 180 180 N-LINKED (GLCNAC. .) (POTENTIAL).
FT NON_TER 269 269
SQ SEQUENCE 269 AA; 29983 MW; 746903F30F95F387 CRC64;

Query Match 12.4%; Score 202; DB 1; Length 269;
Best Local Similarity 28.5%; Pred. No. 8.2e-09;
Matches 67; Conservative 20; Mismatches 122; Indels 26; Gaps 7;

QY 47 LYCAOCPGCTFFVQRCRRDPTGCPGPPRHVYTFQFN---YLERCYCNVLCGEREEA- 102
Db 36 LCCDLCPPGQKLVNDCTEVSKECSGCGEFLSTWNRKYCHEHRYCNPNLGLRIQSEG 95
QY 103 -----RACHATHNRACRCRTGFFAHAGFCLHASCPPGAGVIAPGTPSQTQCPCPGT 157
```

```
Db 96 TLNTDTICVCEGQHTSHT-----CESCTPHSLCLPGFGVKQIATGLDVTCEPCPLGF 150
QY 158 FSASSSSSECOQPHNCTALGLALNVPSSSHDTCTCTGFLPLSTRVPGACECECAVID 217
Db 151 FSNVSAFEKCHRWTSCEKRGVLVEQHVGNKTQDVVC-----GFQSRMRTLWVIVPTMGVLF 206
QY 218 FVAFQDISIKLQRLQLQALEAPEGPGPTPRAGRAALQLKLRRL----TELLGAQD 269
Db 207 AVLLYSACIRNITKKRQ-LRPCTLW----LKGRIPWRRRLIRFFAPATRLSGARD 256

RESULT 13
OX40_MOUSE STANDARD; PRT; 272 AA.
ID OX40_MOUSE
AC P47741;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE OX40L RECEPTOR PRECURSOR (OX40 ANTIGEN).
GN TNFRSF4 OR TXGP1 OR OX40.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C.
RX MEDLINE-94044750; PubMed-8228223;
RA Calderhead D.M., Buhlmann J.E., van den Eertwegh A.J.,
RA Claassen E., Noelle R.J., Fell H.;
RT "Cloning of mouse OX40: a T cell activation marker that may mediate
RT T-B cell interactions.";
RL J. Immunol. 151:5261-5271(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-95255413; PubMed-7737295;
RA Birkeland M.L., Copeland N.G., Gilbert D.J., Jenkins N.A.,
RA Barclay A.N.;
RT "Gene structure and chromosomal localization of the mouse homologue
RT of rat OX40 protein.";
RL Eur. J. Immunol. 25:926-930(1995).
CC -!- FUNCTION: RECEPTOR FOR THE OX40L/GP34 CYTOKINE.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -!-
DR EMBL; Z21674; CAA79772.1; -.
DR EMBL; X85214; CAA59476.1; -.
DR HSSP; P25942; ICDF.
DR MGD; MGI:104512; TXGP1.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; 3.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
KW Receptor; T-cell; Antigen; Glycoprotein; Transmembrane; Repeat;
Signal.
FT SIGNAL 1 19
FT CHAIN 20 272 POTENTIAL.
FT DOMAIN 20 211 OX40L RECEPTOR.
FT TRANSMEM 212 236 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 237 272 POTENTIAL.
FT CYTOPLASMIC (POTENTIAL).
FT DOMAIN 26 165
FT TNFR-CYS 4 X TNFR-CYS.
FT REPEAT 26 61
FT TNFR-CYS 1.
FT REPEAT 62 103
FT TNFR-CYS 2.
FT REPEAT 104 124
FT TNFR-CYS 3 (INCOMPLETE).
FT REPEAT 125 165
FT TNFR-CYS 4.
```

OF AN ALTERNATIVE INITIATION CODON IN THE SAME READING FRAME.

-!- PWM: PHOSPHORYLATED ON SERINE AND TYROSINE RESIDUES.

-!- DISEASE: MOST SPECIFIC HODGKIN'S DISEASE ASSOCIATED ANTIGEN.

-!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.

-!- DATABASE: NAME=PROW; NOTE=CD guide CD30 entry;
WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd30.htm".

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/>) or send an email to license@isb-sib.ch.

EMBL; M83554; AA51947.1; -
DR EMBL; S75768; AAD14188.1; -
DR EMBL; D86042; BAA12973.1; -
DR PIR; A42086; A42086.
DR HSSP; P19438; INCF.
DR MIM; 153243; -
DR INTERPRO; IPR001368; -
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00050; TNFR_NGFR_2; 2.
KW T-cell; Glycoprotein; Transmembrane; Signal; Receptor; Repeat; Phosphorylation; Alternative initiation.
KW SIGNAL 1 18
FT CHAIN 19 595 CD30L RECEPTOR, VARIANT C30V.
FT INIT MET 464 464 FOR VARIANT C30V.
FT DOMAIN 19 379 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 380 407 POTENTIAL.
FT DOMAIN 408 595 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 28 325 6 X TNFR-CYS.
FT REPEAT 28 66 TNFR-CYS 1.
FT REPEAT 68 106 TNFR-CYS 2.
FT REPEAT 107 150 TNFR-CYS 3.
FT REPEAT 205 241 TNFR-CYS 4.
FT REPEAT 243 281 TNFR-CYS 5.
FT REPEAT 282 325 TNFR-CYS 6.
FT DOMAIN 347 377 PRO/SER/THR-RICH.
FT CARBOHYD 101 101 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 595 AA; 63747 MW; 7A407CC78A6E0BC8 CRC64;

Query Match 11.4%; Score 186.5; DB 1; Length 595;
Best Local Similarity 29.9%; Pred. No. 2.8e-07;
Matches 59; Conservative 23; Mismatches 82; Indels 33; Gaps 11;

Qy 15 VLALPALLPVPVAVRGVAET-----PTVWRDAETGERLVCAQCPTGTFVORPCRRDSPTTC 70
Db 10 LLFLGALRAFPQDRPEDTCHGNPSHY---DKAVRRCYRCPMGLFTPTQCPCQ-RPTDC 65
Qy 71 -GPCPPPHYTQFNWYLLERCRYCNVLCGERE-EARACHATHNRCRCRTGFAHAGF--- 125
Db 66 RKQCEPDYLL---DEADRCTAC-VTCRRDLVKTTCANWSSRVCECRGMCSTSAVNS 121
Qy 126 ---CLEHASCPGAGVIATGPTSQNTQCQPCPPGTFSASSSSSEQCQ-----PHRNC 175
Db 122 CARCFHVSVCAGMYKVFPGTAQKNTVCEPAPSGV-SPACASPCNKESKPSGTPQAKPT 180
Qy 176 ALGLALNVPGSSSDTL 192
Db 181 PVS-----PATSSASTM 192

RESULT 15
ID 41BB_MOUSE STANDARD; PRT; 256 AA.
AC P20334;
DT 01-FEB-1991 (Rel. 17, Created)

CARBONYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
CONFLICT 15 15 A -> G (IN REF. 2).
SEQUENCE 272 AA; 30153 MW; 06E7BBA156FD08E CRC64;

Query Match 12.1%; Score 198.5; DB 1; Length 272;
Best Local Similarity 30.3%; Pred. No. 1.5e-08;
Matches 57; Conservative 17; Mismatches 71; Indels 43; Gaps 8;

Qy 15 VLALPALLPVPVAVRGVAETPTVWRDAETGERLVCAQCPTGTFVORPCRRDSPTTCGCP 74
Db 11 LLLALTGLVATRLNCKVHTYP-----SGHK-CCRCQPGHGMVSRCDTRDTLCHPCE 64
Qy 75 PHRYTFQWNY--LERCYCNVLCGEREEARACHATHNRCRCRTGFAHAGCFLHASC 132
Db 65 TGFYNEAVNYDTCKQCTOCNHRSG--SELKQNCPTQDTQVCRCR----- 106
Qy 133 PGAGVIATGPT-PSQNT-----OCQCPCTGFSASSSSPQCPHNRCTALGLALNVP 185
Db 107 -----PGTQPRODSYKLGKVDVCPQPGHF--SPGNQACKPWTNCTLSGKQTRHPA 156
Qy 186 SSSSDTL 193
Db 157 SDSLDAVC 164

RESULT 14
ID CD30_HUMAN STANDARD; PRT; 595 AA.
AC P28908;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE CD30L RECEPTOR PRECURSOR (LYMPHOCYTE ACTIVATION ANTIGEN CD30) (KI-1 ANTIGEN).
GN TNFRSF8 OR CD30.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPHOID;
RX MEDLINE=92154659; PubMed=1310894;
RA Duerkop H., Latza U., Hummel M., Eitelbach F., Seed B., Stein H.;
RT "Molecular cloning and expression of a new member of the nerve growth factor receptor family that is characteristic for Hodgkin's disease.";
RL Cell 68:421-427(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95089787; PubMed=7527901;
RA Jung W., Krueger S., Renner C., Gause A., Sahin U., Trumper L., Pfeundscher H.;
RT "Opposite effects of the CD30 ligand are not due to CD30 mutations: results from cDNA cloning and sequence comparison of the CD30 antigen from different sources.";
RL Mol. Immunol. 31:1329-1334(1994).
RN [3]
RP SEQUENCE FROM N.A. (VARIANT C30V).
RX MEDLINE=96437016; PubMed=8839832;
RA Horie R., Ito K., Tatewaki M., Nagai M., Aizawa S., Watanabe T.; Higashihara M., Ishida T., Inoue J., Takizawa H., Watanabe T.;
RT "A variant CD30 protein lacking extracellular and transmembrane domain is induced in HL-60 by tetradecanoylphorbol acetate and is expressed in alveolar macrophages.";
RL Blood 88:2422-2432(1996).
CC -!- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD30L. MAY PLAY A ROLE IN THE REGULATION OF CELLULAR GROWTH AND TRANSFORMATION OF ACTIVATED LYMPHOBLASTS. REGULATES GENE EXPRESSION THROUGH ACTIVATION OF NFkB.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- ALTERNATIVE PRODUCTS: A SHORTER CYTOPLASMIC FORM C30V WHICH IS ONLY EXPRESSED IN ALVEOLAR MACROPHAGES IS PRODUCED BY THE USE

us-09-518-931-2.rsp

Wed Jan 31 13:43:51 2001

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:39:44 ; Search time 254.27 Seconds
(without alignments)
138.288 Million cell updates/sec

Title: US-09-518-931-2

Perfect score: 1634

Sequence: 1 MRALGPGLSLCLVLALPA.....RVARMPGLSRVRFRLPVH 300

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_15:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phase:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1634	100.0	300	4	O95407
2	444	27.2	401	4	O00300
3	440.5	27.0	372	4	O9UHP4
4	425.5	26.0	401	11	O08727
5	424.5	26.0	401	11	O08712
6	395	24.2	302	13	O9PUS0
7	340.5	20.8	439	4	Q16042
8	333.5	20.4	459	11	O62327
9	327	20.0	482	11	O88734
10	287	17.6	655	4	O75509
11	282.5	17.3	348	12	O57277
12	280.5	17.2	348	12	O57103
13	280.5	17.2	348	12	O57108
14	276	16.9	349	12	O57100
15	274	16.8	349	12	O57291
16	274	16.8	349	12	O57101
17	274	16.8	349	12	O57102
18	273	16.7	349	12	O57099
19	265	16.2	350	12	O57116

20	264.5	16.2	355	12	Q85308
21	262.5	16.1	349	12	O57110
22	262.5	16.1	349	12	O57111
23	262.5	16.1	349	12	Q89118
24	262.5	16.1	349	12	Q89098
25	262	16.0	348	12	O57112
26	262	16.0	348	12	Q85407
27	261.5	16.0	349	12	O57284
28	261.5	16.0	349	12	O57098
29	259.5	15.9	349	12	O57305
30	259.5	15.9	360	12	O57118
31	258.5	15.8	351	12	O57117
32	258.5	15.8	351	12	O73559
33	257.5	15.8	326	12	O57120
34	257.5	15.8	349	12	O57097
35	254.5	15.6	326	12	O57122
36	254.5	15.6	349	12	O57109
37	253.5	15.5	347	12	O57115
38	251.5	15.4	347	12	O57119
39	250.5	15.3	351	12	O57121
40	249	15.2	283	6	Q9XS28
41	246	15.1	283	4	Q92956
42	246	15.1	283	4	Q9UM65
43	245.5	15.0	350	12	O57123
44	239	14.6	616	4	Q9Y606
45	233.5	14.3	625	11	O35305

ALIGNMENTS

RESULT 1

O95407	AC	O95407	PRELIMINARY;	PRT;	300 AA.
DT	01-MAY-1999	(TRENBLrel. 10, Created)			
DT	01-MAY-1999	(TRENBLrel. 10, Last sequence update)			
DT	01-OCT-2000	(TRENBLrel. 15, Last annotation update)			
DE	DECOY RECEPTOR 3 (M68)	(M68C) (M68E).			
GN	DCR3 OR TR6.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=99087326; PubMed=9872321;				
RA	Pitti R.M., Marsters S.A., Lawrence D.A., Roy M., Kischkel F.C.,				
RA	Dowl P., Huang A., Donahue C.J., Sherwood S.W., Baldwin D.T.,				
RA	Godowski P.J., Wood W.I., Gurney A.L., Hillan K.J., Cohen R.L.,				
RA	Goddard A.D., Botstein D., Ashkenazi A.				
RT	"Genomic amplification of a decoy receptor for Fas ligand in lung and				
RT	colon cancer."				
RL	Nature 396:699-703(1998).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=BLOOD.				
RX	MEDLINE=99253915; PubMed=10318773;				
RA	Yu K.Y., Kwon B., Ni J., Zhai Y., Ebner R., Kwon B.S.;				
RT	"A newly identified member of tumor necrosis factor receptor				
RT	superfamily (TR6) suppresses LIGHT-mediated apoptosis."				
RN	J. Biol. Chem. 274:13733-13736(1999).				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=PANCREAS;				
RX	MEDLINE=20122600; PubMed=10655513;				
RA	Bai C., Connolly B., Metzker M.L., Hilliard C.A., Liu X., Sandig V.,				
RA	Soderman A., Galloway S.M., Liu Q., Austin C.P., Caskey C.T.;				
RT	"Overexpression of M68/DCR3 in human gastrointestinal tract tumors				
RT	independent of gene amplification and its location in a four-gene				
RT	cluster."				
RL	Proc. Natl. Acad. Sci. U.S.A. 97:1230-1235(2000).				
DR	EMBL; AF104419; AAD03056.1; "				

Q85308	cowpox viru
O57110	variola vir
O57111	variola vir
Q89118	variola vir
Q89098	variola vir
O57112	variola vir
Q85407	variola vir
O57284	camelpox vi
O57098	camelpox vi
O57305	cowpox viru
O57118	cowpox viru
O57117	cowpox viru
O73559	cowpox viru
O57120	cowpox viru
O57097	camelpox vi
O57122	cowpox viru
O57109	variola vir
O57115	cowpox viru
O57119	cowpox viru
O57121	cowpox viru
Q9XS28	carcopithec
Q92956	homo sapien
Q9UM65	homo sapien
O57123	cowpox viru
Q9Y606	homo sapien
O35305	mus musculu


```
Db 61 VCAPCPDHYDTHSWHTSDECLYSPVCKELQYVKQECNRTHNRVCECKEGRYLEIEFCLK 120
Qy 129 HASCPGAGVIAPTPSONTCQCPCTCFSSSSSSQCOQHNRCTALGLALNVPGSSS 188
Db 121 HRSPPGFGVQAGTPERTVCKKCPDGFNFSSKAPCKRHTNCSVFGLLLTQKGNAT 180
Qy 189 HDTLCTSCGFPGLSTRVPGAEE--CERAVIDF 218
Db 181 HDNI---CSGNSSESTQKCGIDVTLCCEAERFF 209

RESULT 3
Q9UHP4 PRELIMINARY; PRT; 372 AA.
AC Q9UHP4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE OSTEOPROTEGERIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN SEQUENCE FROM N.A.
RP TISSUE=PLACENTA;
RA He Z.-Y., Yang G.-Z., Zhang W.-J., Wu X.-F.;
RT "Cloning and Expression of Osteoprotegerin from Homo sapiens.";
RL Sheng Wu Hua Hsueh Yu Sheng Wu Li Hsueh Pao 31:680-684(1999).
DR EMBL; AF134187; AAF20168.1; -.
DR HSSP; P25942; 1CDF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PSS0050; TNFR_NGFR_2; 2.
FT NON_TER 1
SQ SEQUENCE 372 AA; 42758 MW; F02527A5CD01CCD3 CRC64;

Query Match 27.08; Score 440.5; DB 4; Length 372;
Best Local Similarity 41.28; Pred. No. 1.9e-31;
Matches 77; Conservative 31; Mismatches 74; Indels 5; Gaps 2;

Qy 34 PTYPWRDAETGERLVCAOCPGTFVQRCRDSPTTCGCPPRHYTFQWNYLERCYCNV 93
Db 5 PKYLHYDEETSHQLLCKCPGTYLKQCTAKWTVCAPCPDHYTDSWHTSDECLYCS 64
Qy 94 LGEREEEARACHATHNRACRRTGFFAHAGFCLHSHASCPGAGVIAPGTPSONTCQCP 153
Db 65 VCKELQYVKQECNRTHNRVCECKEGRYLEIEFCLKHRSCPPGFGVQAGTPERNVCKRC 124
Qy 154 PPGTFSSASSSSQCOQHNRCTALGLALNVPGSSSHDILCTSCGFPPLSTRVPGAEE--C 211
Db 125 PDGFFSNSTSKAPCKRHTNCSVFGLLLTQKGNATHDNI---CSGNSSESTQKCGIDVTLC 181
Qy 212 ERAVIDF 218
Db 182 EEAERFF 188

RESULT 4
ID O08727 PRELIMINARY; PRT; 401 AA.
AC O08727;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)
DE (OCIF).
GN TNFRSF11B OR OPG.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
```

```
OX NCBI_TaxID=10116;
RN SEQUENCE FROM N.A.
RP TISSUE=INTESTINE;
RX MEDLINE=97262071; PubMed=9108485;
RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,
RA Luethy R., Nguyen H.Q., Woodson S., Bennett L., Boone T., Shimamoto G.,
RA Derose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,
RA Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,
RA Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,
RA Suggs S., Boyle W.J.;
RT "Osteoprotegerin: a novel secreted protein involved in the regulation
RL Cell 89:309-319(1997).
CC -!- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES
CC OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY
CC SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO
CC OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN
CC STROMAL CELLS AND OSTEOCLAST PROGENITORS.
CC -!- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM) (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL; U94330; AAB53707.1; -.
DR HSSP; P25942; 1CDF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PSS0050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -.
KW Glycoprotein; Repeat; Cytokine; Signal.
FT SIGNAL 1 21
FT CHAIN 22 401
FT DOMAIN 23 201
FT REPEAT 23 63
FT REPEAT 64 106
FT REPEAT 107 143
FT REPEAT 144 201
FT DOMAIN 306 365
FT DISULFID 41 54
FT DISULFID 44 62
FT DISULFID 65 80
FT DISULFID 83 97
FT DISULFID 87 105
FT DISULFID 118 142
FT DISULFID 145 160
FT CARBOHYD 98 98
FT CARBOHYD 165 165
FT CARBOHYD 178 178
FT CARBOHYD 289 289
SQ SEQUENCE 401 AA; 46192 MW; FEC6A31F1D4E573A CRC64;

Query Match 26.08; Score 425.5; DB 11; Length 401;
Best Local Similarity 39.58; Pred. No. 4.5e-30;
Matches 81; Conservative 33; Mismatches 86; Indels 5; Gaps 2;

Qy 34 PTYPWRDAETGERLVCAOCPGTFVQRCRDSPTTCGCPPRHYTFQWNYLERCYCNV 93
Db 26 PKYLHYDEETSHQLLCKCPGTYLKQCTAKWTVCAPCPDHYTDSWHTSDECLYCS 85
Qy 94 LGEREEEARACHATHNRACRRTGFFAHAGFCLHSHASCPGAGVIAPGTPSONTCQCP 153
Db 86 VCKELQTVKQECNRTHNRVCECKEGRYLEIEFCLKHRSCPPGFGVQAGTPERNVCKRC 145
Qy 154 PPGTFSSASSSSQCOQHNRCTALGLALNVPGSSSHDILCTSCGFPPLSTRVPGAEE--C 211
Db 146 PDGFFSGSTSSKAPCKRHTNCSVFGLLLTQKGNATHDNI---CSGNSSESTQKCGIDVTLC 202
Qy 212 ERAVIDFVAFODISIKRLQRLQAL 236
Db 203 EEAERFFAVPTKIIPNWLVLVDSL 227
```

RESULT	5
08712	PRELIMINARY; PRT: 401 AA.
O08712; 070202;	
I 01-JUL-1997 (T-EMBLrel. 04, Created)	
T 01-JUL-1997 (T-EMBLrel. 04, Last sequence update)	
F 01-JUL-1997 (T-EMBLrel. 15, Last annotation update)	
E 01-OCT-2000 (T-EMBLrel. 15, Last annotation update)	
E OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR) (OCIF).	
E TNFRSF11B OR OPG.	
S Mus musculus (Mouse).	
C Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
C Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
X NCBI_TaxId=10090;	
[1]	
SEQUENCE FROM N.A.	
P STRAIN-BALB/C; TISSUE-KIDNEY;	
X MEDLINE=97262071; PubMed=9108485;	
X Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,	
A Leuchy R., Nguyen H.Q., Woodson S., Bennett L., Boone T., Shimamoto G.,	
A Derose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,	
A Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,	
A Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,	
A Sugis S., Boyle W.J.;	
T "Osteoprotegerin: a novel secreted protein involved in the regulation	
T of bone density.";	
L Cell 89:309-319(1997).	
[2]	
SEQUENCE FROM N.A.	
P STRAIN-129/Ola. AND NIH SWISS:	
X MEDLINE=9838527; PubMed=9714833;	
X Mizuno A., Murakami A., Nakagawa N., Yasuda E., Morinaga T.,	
A Higashio K.;	
V "Structure of the mouse osteoclastogenesis inhibitory factor (OCIF)	
V gene and its expression in embryogenesis.";	
CC Gene 215:339-343(1998).	
-1 FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES	
CC OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY	
CC SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO	
CC OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN	
CC STROMAL CELLS AND OSTEOCLAST PROGENITORS.	
-1 SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM).	
-1 SUBCELLULAR LOCATION: EXTRACELLULAR.	
-1 TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN LIVER, LUNG,	
CC BRAIN, HEART, KIDNEY, STOMACH, INTESTINE, SKIN, CALVARIA AND	
CC PLACENTA. NOT DETECTED IN SPLEEN.	
-1 DEVELOPMENTAL STAGE: IN THE EMBRYO, HIGH LEVELS ARE DETECTED AT	
CC DAY 7. AT DAY 11, EXPRESSION DECREASES AND THEN INCREASES FROM DAY	
CC 15 TO DAY 17.	
-1 SIMILARITY: CONTAINS A LA-NGF/TNFR-TYPE CYSTEINE-RICH REGION.	
CC EMBL; U94331; AAB53708.1; ..	
DR EMBL; AB013898; BAA28269.1; ..	
DR EMBL; AB013903; BAA33388.1; ..	
DR EMBL; AB013899; BAA33388.1; JOINED.	
DR EMBL; AB013900; BAA33388.1; JOINED.	
DR EMBL; AB013901; BAA33388.1; JOINED.	
DR EMBL; AB013902; BAA33388.1; JOINED.	
DR HSSP; P25942; 1CDF.	
DR MGD; MG1_109587; Opg.	
DR INTERPRO; IPR000488; ..	
DR INTERPRO; IPR001388; ..	
DR PFAM; PF00020; TNFR_C6; 3.	
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.	
DR PROSITE; PS50017; DEATH_DOMAIN; 1.	
DR PROSITE; PS50050; TNFR_NGFR_2; 2.	
DR PRODOM; PD000771; -; 1.	
DR Glycoprotein; Repeat; Cytokine; Signal.	
KW SIGNAL 1 21	
FT CHAIN 22 401	OSTEOPROTEGERIN.
FT DOMAIN 23 201	4 X TNFR-CYS.
FT REPEAT 24 63	TNFR-CYS 1.
FT REPEAT 54 106	TNFR-CYS 2.
FT	

DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS50050; TNFR_NGFR_2; 1.
KW Receptor.
SQ SEQUENCE 302 AA; 34037 MW; E44C73477F05C3DF CRC64;

Query Match 24.2%; Score 395; DB 13; Length 302;

Best Local Similarity 34.3%; Pred. No. 1.7e-27;

Matches 82; Conservative 44; Mismatches 101; Indels 12; Gaps 7;

Qy 35 TYPHRAETGERLVCAQCPGPTFFVQRCRRDSPITCGCPPHRYTQFWNYLERCRYCNVL 94
Db 22 TFKNDRIYSGLSIVDCRPPGTYLRAPCSAMRKSDCAECNGAYTEFWNHISKLRCS-M 80
Qy 95 CGEEREERACHATHNRACRCRTGFEAHGF--CLEHASCPPGAGVAPGTPSNTOCOP 152
Db 81 CAENOVVKQECSPNSNCECEKEGYFNKKYEAKIKHKECPGPGYANTGTPHODTECV 140
Qy 153 CPPGTFSASSSSQOCQPHRNCTALGLALNVPGSSSHDTLCTSGTGFPLSTRVPGAEC 212
Db 141 COAGFYSEVSAKATCLAQSNCKVGLRVLKGQDWHNTLCASC--YDLKTR-DGAEYLH 197
Qy 213 RAVIDFV--AFQDISIKRLQLLQALEAPEGWGPTPRAGRAALQKLRRLTELIGAOD 269
Db 198 EILPTFFIQLHQTGMIKMRRL--AMRLPQEGGKKPLIG--AVMKRNRRLHDFMNSWD 252

RESULT 7

Q16042 ID Q16042 PRELIMINARY; PRT; 439 AA.
AC Q16042;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-JAN-1999 (TREMBlrel. 02, Last sequence update)
DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91370690; PubMed=1966549;
RA Dembic Z., Loetscher H., Gubler U., Pan Y.C., Lahm H.W., Gentz R.,
RA Brockhaus M., Lesslauer W.
RT "Two human TNF receptors have similar extracellular, but distinct
RT intracellular, domain sequences".
RL Cytochrome 2:231-237(1990).
DR EMBL; S63368; AAB19824.1; -.
DR HSSP; P25942; 1CDF.
DR INTERPRO: IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 3.
DR PRODOM; PD000771; -.
SQ SEQUENCE 439 AA; 46090 MW; FEBCB329CC67FF6 CRC64;

Query Match 20.8%; Score 340.5; DB 4; Length 439;

Best Local Similarity 29.7%; Pred. No. 1.8e-22;

Matches 89; Conservative 43; Mismatches 115; Indels 53; Gaps 11;

Qy 30 VAETPTYPHRAETGE-----RLVCAQCPGPTFFVQRCRRDSPITCGCPPHR 77
Db 5 VAFTPYAP-----EPGSTCRLEIYYDQTQMCCSKSPGQHKVFTKTSYDVCDSDEST 60
Qy 78 YTFQWNYLERCRYCNVLCGEREEERACHATHNRACRCRTGFFAHAG-----FCLHAS 131
Db 61 YTOLNNWYPECLSCGSRSSQDVETQACTREQNRICTRCPGWYCALSKQECRLCAPLRK 120
Qy 132 CPPGAGVITAPGTPSNTOCQCPGPTGTFSSASSSSQOCQPHRNCTALGLALNVPGSSSHDT 191
Db 121 CRPGFGVAPGPTGTSDDVVKCKPCAPGPTFSNTTSSTDICRPHQICNVVA----IPGNASMDA 176

Qy 192 LCTCTGFPPLSTRVPGAECERAVIDFVAFQDISIKRLQLLQALEAPE-----GWGTP 246
Db 177 VCTSTS--PTRSMAPGAVHLPQPV-----STRSOHTQPTPEPSTAPSTFLLPMGSP 227
Qy 247 RA----GRAALQKLRRLTELGAQDQALLVRLQAL----RVARMP-GLERSVREREL 298
Db 228 PAEGSTGDFALPVLGVLTAL-----GLLIIGVNVVIMTQVKKKPLCLQREAKVPHLP 282

RESULT 8

Q62327 ID Q62327 PRELIMINARY; PRT; 459 AA.
AC Q62327;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR 2 MRNA (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NOD;
RX MEDLINE=95178848; PubMed=7873884;
RA Powell E.E., Wicker L.S., Peterson L.B., Todd J.A.;
RT "Allelic variation of the type 2 tumor necrosis factor receptor
RT gene".
RL Mamm. Genome 5:726-727(1994).
DR EMBL; X76401; CAA53981.1; -.
DR HSSP; P19438; 1NCF.
DR INTERPRO: IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 3.
FT NON_TER 1
FT VARIANT 87 87 S->T.
FT VARIANT 93 93 T->I.
FT VARIANT 268 268 F->I.
FT VARIANT 345 345 S->F.
FT VARIANT 421 421 Y->C.
SQ SEQUENCE 459 AA; 48686 MW; 6C51D2CF1C4626DF CRC64;

Query Match 20.4%; Score 333.5; DB 11; Length 459;

Best Local Similarity 29.7%; Pred. No. 7.8e-22;

Matches 81; Conservative 43; Mismatches 110; Indels 39; Gaps 9;

Qy 46 RLVCAQCPGPTFFVQRCRRDSPITCGCPPHRYTQFWNYLERCRYCNVLCGEREEERAC 105
Db 37 OMCCAKCPGQYVKHFCNKTSDTVCADCEASMYTQVWNOFRTCLSCSSCSTDQVETRAC 96
Qy 106 HATHNRACRCRTGTF-----AHAGF---CLEHASCPPGAGVITAPGTPSNTOCQCPGPT 158
Db 97 TKQQRVCAECAGRYCALKTHSGSCRCMRLSKCPGFGVASSRAPNGVNLKACAPGTF 156
Qy 159 SASSSSQOCQPHRNCTALGLALNVPGSSSHDTLCT-----SCTGPPLSTRVPGAECERA 214
Db 157 SDTTSSTVCRPHRICSTLA----IPGNASTDAVCAPESTLSAIPRTLYSQQEPTRSQ 212
Qy 215 VIDFVAFQDISIKRLQLLQALEAPEGWGPTP-----RAGRAALQKLRRLTELIGAOD 269
Db 213 PLD-----QEPGSPQPSILTSL-----GSTPIIBQSTKGGISLPIGLIVGVTSL----- 257
Qy 270 GALLVRLQAL----RVARMPGLERSVREREL 298
Db 258 GLLMGLVNCFTLVQKKKPSCLQORDAKVPHVP 290

RESULT 9

O88734 ID O88734 PRELIMINARY; PRT; 482 AA.
AC O88734;

Identification and functional characterization of DR6, a novel death domain-containing TNF receptor.
Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
[2]

RT 01-NOV-1998 (TReMBLrel. 08, Created)
RT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
RL 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
RN P80 TNF-ALPHA RECEPTOR.
GN TNFR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RA Hurler B., Segade F., Rodriguez R., Ramos S.S., Lazo P.S.;
RA "The Mouse Tumor Necrosis Factor Receptor Gene:Genomic Structure and
RT Characterization of the two Transcripts.";
RL Genomics 0:0-0(0).
DR EMBL; Y14619; CAA74969.1; JOINED.
DR EMBL; Y14620; CAA74969.1; JOINED.
DR EMBL; Y14621; CAA74969.1; JOINED.
DR EMBL; Y14622; CAA74969.1; JOINED.
DR EMBL; Y14623; CAA74969.1; JOINED.
DR EMBL; Y14679; CAA74969.1; JOINED.
DR HSSP; P19438; INCF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_c6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00050; TNFR_NGFR_2; 3.
DR PRODOM; PD000771; -. 1.
SQ SEQUENCE 482 AA; 51106 MW; F6C15046B48FF83C CRC64;

Query Match 20.08; Score 327; DB 11; Length 482;
Best Local Similarity 29.3%; Pred. No. 3.1e-21;
Matches 82; Conservative 43; Mismatches 109; Indels 46; Gaps 10;
QY 46 RLVCACPPGTFVORPCRR-----DSPTTCGCPGPRHNTQFNLYLRCRYCNVL 98
DB 52 QMCCAKCPGQYVHFCNKTSVTCADSDTVACDCEASMTQVNNQFRTCLSCSSSSTD 111
QY 99 EEEERACHATHNRCRCRTGTF-----AHAGF---CLEHACPPGAGVIAPCTPSQTCQ 151
DB 112 QVETRACKQNRVCACGAGRYCALKTHSGSCRCQMRSLKCGPGFGVASSRAPNGVLCK 171
QY 152 PCPPGTFSSSSSSQOPHRNCTALGLALNVPGSSSHDTLCT---SCTGFPPLSTRVPG 207
DB 172 ACAPGTFSTSTSDVCRPHRCISLA---IPGNASTDAVCAPESTLSAIPRTLVSQ 227
QY 208 AEECERAVIDFAVDISTKRLQRLQALEAPEGHWGTP-----RAGRAALQLRLRLT 262
DB 228 PEPTRSQPLD---QRPGPSQTPSITSL-----GSTPIEQSTKGISLPIGLVGVT 277
QY 263 ELLGAQDQALLVLLQAL----RVARMPLGLRSVRERFLP 298
DB 278 SL-----GLMLGLVNCVFILVORKKPKSLQDKAVPHVP 312

RESULT 10
ID 075509 PRELIMINARY; PRT; 655 AA.
AC 075509;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE TNFR-RELATED DEATH RECEPTOR-6 (DJ181J13.1) (DR6 (TNFR-RELATED DEATH RECEPTOR-6)).
GN DR6 OR DJ181J13.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RA Pan G., Bauer J.H., Haridas V., Wang S., Liu D., Ni J., Yu G.,
RA Vincenz C., Aggarwal B.B., Dixit V.M.;

Query Match 17.6%; Score 287; DB 4; Length 655;
Best Local Similarity 34.3%; Pred. No. 1.5e-17;
Matches 58; Conservative 23; Mismatches 88; Indels 0; Gaps 0;
QY 35 TYPWRDAETGERLVCACQCPGTFVORPCRRDSPTTCGPPRHNTQFNLYLRCRYCNVL 94
DB 53 TYRHVDRTAGVLTCDKCPAGTYVSEHCTNTSLRVCSSCPVGTFRHENGIEKHCDSQP 112
QY 95 CGEEERACHATHNRCRCRTGTFFAHAGFCLEHACPPGAGVIAPCTPSQTCQCP 154
DB 113 CPWPMIEKLPCALTDRCCTCPPGFMFQSNATCAPHTVCPVGWGRKKGTETEDVRCQCA 172
QY 155 PGTFSASSSSQOPHRNCTALGLALNVPGSSSHDTLCTSCTGFPPLST 203
DB 173 RGTFSDFVPSVVMKCRAYTDCLSNLVVWIKPGTKETDNVCGTLPSFSST 221

RESULT 11
ID 057277 PRELIMINARY; PRT; 348 AA.
AC 057277;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN CRMB.
OS Monkeypox virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus
OX NCBI_TaxID=10244;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=ZAIRE-1996(96-17), ZAIRE-1996 (96-16);
RA Loparev V.N., Parsons J.M., Esposito J.J.;
RA Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RL EMBL; U88543; AAB94378.1; -.
DR EMBL; U87841; AAB94358.1; -.
DR HSSP; P25942; ICDF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_c6; 2.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -. 1.
SQ SEQUENCE 348 AA; 38212 MW; 54019521556C2D8F CRC64;

Query Match 17.3%; Score 282.5; DB 12; Length 348;
Best Local Similarity 30.9%; Pred. No. 2e-17;
Matches 64; Conservative 33; Mismatches 97; Indels 13; Gaps 3;
QY 9 LSLCLVLALPALLPVPVAVRGVAETPTYPWRDAETGERLVCACQCPGTFVORPCRRDSPT 68

```
Db 10 LFLSCIIINGRDLAPHSNGKCKDNEYRSR-----LCCLSCPPGYASRLCDSKNT 63
Qy 69 TCGPCPPRHYYTFQWNYLERCRYCNVLCGEREEARACHATHNRACRGTGFF-----AH 122
Db 64 OCTPCGSDTFTSHNNHLQACLSCNGCRDSNQVETRSCNTHNRICECSPGYYCLLKSGSSG 123
Qy 123 AGFCLEHASCPPGAGVIAPGTPSQNTQCPGPGTFSASSSEOCOPHRNCTALGLALN 182
Db 124 CRTICSKTKCGIGYV-SGYTSTGVDVCSGPCGPGYSHTVSSTDCKEPTVTSNTFNYIDVE 182
Qy 183 VPGSSSHDTLCTSCGTGFLPSTRVPGAE 209
Db 183 INLYPVNDTSCRTTTTGLSESISTSE 209

RESULT 12
O57108
ID O57103 PRELIMINARY; PRT; 348 AA.
AC O57103;
DT 01-JUN-1998 (TremBrel. 06, Created)
DT 01-JUN-1998 (TremBrel. 06, Last sequence update)
DT 01-OCT-2000 (TremBrel. 15, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN CRMB.
OS Monkeypox virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10244;
RN [1]
RC STRAIN=Zaire-1979 (79-0005);
RA Loparev V.N., Parsons J.M., Esposito J.J.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U87847; AAB94364.1; -.
DR HSSP; P25942; 1CDF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 2.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -.
SQ SEQUENCE 348 AA; 38184 MW; 34A5E668B27907B5 CRC64;
```

```
Query Match 17.2%; Score 280.5; DB 12; Length 348;
Best Local Similarity 30.4%; Pred. No. 3e-17;
Matches 63; Conservative 34; Mismatches 97; Indels 13; Gaps 3;
```

```
Qy 9 LSLCLVLALPALLPVPVAVRGVAETPTYPWRDAETGERLVCAQCPCPGTFVQPCRRDSPT 68
Db 10 LFLSCIIINGRDLAPHSNGKCKDNEYRSR-----LCCLSCPPGYASRLCDSKNT 63
Qy 69 TCGPCPPRHYYTFQWNYLERCRYCNVLCGEREEARACHATHNRACRGTGFF-----AH 122
Db 64 OCTPCGSDTFTSHNNHLQACLSCNGCRDSNQVETRSCNTHNRICECSPGYYCLLKSGSSG 123
Qy 123 AGFCLEHASCPPGAGVIAPGTPSQNTQCPGPGTFSASSSEOCOPHRNCTALGLALN 182
Db 124 CRTICSKTKCGIGYV-SGYTSTGVDVCSGPCGPGYSHTVSSTDCKEPTVTSNTFNYIDVE 182
Qy 183 VPGSSSHDTLCTSCGTGFLPSTRVPGAE 209
Db 183 INLYPVNDTSCRTTTTGLSESISTSE 209
```

```
RESULT 13
O57108
ID O57108 PRELIMINARY; PRT; 348 AA.
AC O57108;
```

```
DT 01-JUN-1998 (TremBrel. 06, Created)
DT 01-JUN-1998 (TremBrel. 06, Last sequence update)
DT 01-OCT-2000 (TremBrel. 15, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
```

```
GN CRMB.
OS Monkeypox virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10244;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=Zaire-1970 (CONGO-8);
RA Loparev V.N., Parsons J.M., Esposito J.J.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U88142; AAB94367.1; -.
DR HSSP; P25942; 1CDF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 2.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -.
SQ SEQUENCE 348 AA; 38212 MW; E555979057DEC91F CRC64;
```

```
Query Match 17.2%; Score 280.5; DB 12; Length 348;
Best Local Similarity 30.4%; Pred. No. 3e-17;
Matches 63; Conservative 34; Mismatches 97; Indels 13; Gaps 3;
```

```
Qy 9 LSLCLVLALPALLPVPVAVRGVAETPTYPWRDAETGERLVCAQCPCPGTFVQPCRRDSPT 68
Db 10 LFLSCIIINGRDLAPHSNGKCKDNEYRSR-----LCCLSCPPGYASRLCDSKNT 63
Qy 69 TCGPCPPRHYYTFQWNYLERCRYCNVLCGEREEARACHATHNRACRGTGFF-----AH 122
Db 64 OCTPCGSDTFTSHNNHLQACLSCNGCRDSNQVETRSCNTHNRICECSPGYYCLLKSGSSG 123
Qy 123 AGFCLEHASCPPGAGVIAPGTPSQNTQCPGPGTFSASSSEOCOPHRNCTALGLALN 182
Db 124 CRTICSKTKCGIGYV-SGYTSTGVDVCSGPCGPGYSHTVSSTDCKEPTVTSNTFNYIDVE 182
Qy 183 VPGSSSHDTLCTSCGTGFLPSTRVPGAE 209
Db 183 INLYPVNDTSCRTTTTGLSESISTSE 209
```

```
RESULT 14
O57100
ID O57100 PRELIMINARY; PRT; 349 AA.
AC O57100;
```

```
DT 01-JUN-1998 (TremBrel. 06, Created)
DT 01-JUN-1998 (TremBrel. 06, Last sequence update)
DT 01-OCT-2000 (TremBrel. 15, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
```

```
GN CRMB.
OS Monkeypox virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10244;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=NIGERIA-1971 (71-0082);
RA Loparev V.N., Parsons J.M., Esposito J.J.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U87844; AAB94361.1; -.
DR HSSP; P25942; 1CDF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 2.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -.
SQ SEQUENCE 349 AA; 38239 MW; DF6C280D478F2422 CRC64;
```

```
Query Match 16.9%; Score 276; DB 12; Length 349;
Best Local Similarity 30.0%; Pred. No. 7.6e-17;
Matches 63; Conservative 34; Mismatches 95; Indels 18; Gaps 5;
```

```

RESULT 15
057291 PRELIMINARY; PRT; 349 AA.
057291;
01-JUN-1998 (TEmBLrel. 06, Created)
01-JUN-1998 (TEmBLrel. 06, Last sequence update)
01-OCT-2000 (TEmBLrel. 15, Last annotation update)
TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
CRMB.
OS Monkeypox virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
NCBI_TaxID=10244;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-VARIOUS STRAINS;
RC Loparev V.N., Parsons J.M., Esposito J.J.;
RA Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RL EMBL; U88144; AAB94369.1; -
DR EMBL; U87842; AAB94359.1; -
DR EMBL; U87994; AAB94365.1; -
DR EMBL; U87995; AAB94366.1; -
DR EMBL; U88143; AAB94368.1; -
DR HSP; P25942; ICDF.
DR INTERPRO; IPR001368; -
DR PFAM; PF00020; TNFR_c6; 2.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -; 1.
DR PRODOM; 349 AA; 38295 MW; CBD2C949ED2B8E7C CRC64;
SO SEQUENCE

```

	Query Match	16.8%;	Score 274;	DB 12;	Length 349;	
	Best Local Similarity	30.0%;	Prod. No. 1.1e-16;			
	Matches 63;	Conservative 34;	Mismatches 95;	Indels 18;	Gaps 5;	
QY	9	LSLCLVLALPALLVPVAVRGAFTPTYPWRDAETGERLVCAQCPCPGTVQVPCRRDSPT	68			
DB	10	LFLSCIIINGRIDIAHPASNGKCKDNEYRSN-----LCCLSCPSTVASRLCDSKNT	63			
QY	69	TCGCPPPRHYTFQWNYLERCYNCVLCGBREERACHATHNACRCRTGFF-----AH	122			
DB	64	QCTPCGSDTFTSHNHLQALCISLNGRCDNQVETRCNTTHNRCIBCSPGYCYLLKGASG	123			
QY	123	AGFCLEHASCPCGGAVTAGPFSQNTQCPCPGTTSASSSSSEOCQ---HRCNTALGL	179			
DB	124	CRTCLSKTKCGIGYV--SGYTSSTDGVICSPCGPGCTYSHTVSTDKCEPVVTSNTFNYIDV	182			
QY	180	ALNVPGSSSHDFTLCTSCGTGFFLSTRVFGAE	209			
DB	183	EINL--YPVNDTSCTRTTTTGLSESISTSE	210			

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:45:20 ; Search time 183.67 Seconds

(without alignments)
2.979 Million cell updates/sec

Title: US-09-518-931-2_COPY_31_46

Perfect score: 92

Sequence: 1 AETPTYPWRDAETGER 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq_36: *
- 1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT:*
 - 2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT:*
 - 3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT:*
 - 4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT:*
 - 5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT:*
 - 6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT:*
 - 7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT:*
 - 8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT:*
 - 9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT:*
 - 10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT:*
 - 11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT:*
 - 12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT:*
 - 13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT:*
 - 14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT:*
 - 15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT:*
 - 16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT:*
 - 17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT:*
 - 18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT:*
 - 19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT:*
 - 20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT:*
 - 21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	92	100.0	170	19 W63623	Human tumour necro
2	92	100.0	245	20 Y28449	A human tumour nec
3	92	100.0	271	20 Y42184	Human mFLINT #1 pr
4	92	100.0	273	20 Y42185	Human mFLINT #2 pr
5	92	100.0	300	19 W66102	Amino acid sequenc
6	92	100.0	300	19 W63622	Human tumour necro
7	92	100.0	300	20 Y03099	Human lung TNF-rec
8	92	100.0	300	20 Y42182	Human FLINT #1 pro
9	92	100.0	300	20 Y17479	Mammalian tumour n
10	92	100.0	300	20 Y06817	Human DcR3 polypep
11	92	100.0	300	20 W97749	Human tumour necro
12	92	100.0	300	20 W95082	Orphan receptor (H

13	92	100.0	300	21 Y77458	Human TNF receptor
14	92	100.0	302	20 Y42183	Human FLINT #2 pro
15	49	53.3	355	19 W40382	S. glaucescens acb
16	47	51.1	775	19 W58582	Kojibiose phosphor
17	45	48.9	106	19 Y20986	Human glial fibrill
18	45	48.9	7257	21 Y58576	Sorangium cellulos
19	43.5	47.3	1421	21 Y58573	Sorangium cellulos
20	41	44.6	199	20 Y14965	Partial amino acid
21	41	44.6	397	5 P40063	Sequence of aminot
22	41	44.6	397	9 P80686	Protein encoded by
23	41	44.6	797	17 W06821	Infectious bursal
24	40	43.5	122	16 R66628	Hepatitis C virus
25	40	43.5	122	20 Y01619	Protein encoded by
26	40	43.5	122	20 W30593	Hepatitis C virus
27	40	43.5	131	13 W41080	Non-A non-B hepati
28	40	43.5	268	14 R34007	HCV J1.1 E2/NS1 HV
29	40	43.5	268	14 R34008	HCV J1.2 E2/NS1 HV
30	40	43.5	269	14 R34001	JH-1 E2/NS1 protei
31	40	43.5	347	19 W37127	Hepatitis C virus
32	40	43.5	352	20 Y37260	Protein which is s
33	40	43.5	395	20 Y01620	Protein encoded by
34	40	43.5	402	14 R34433	Sequence of glycop
35	40	43.5	416	16 R66629	Hepatitis C virus
36	40	43.5	416	20 W30598	Hepatitis C virus
37	40	43.5	539	19 W37128	Hepatitis C virus
38	40	43.5	3010	14 R34580	Human hepatitis C
39	40	43.5	3011	14 R34468	Encoded by full-le
40	40	43.5	3014	14 R35207	Hepatitis C virus
41	40	43.5	3014	15 R54099	NANBHV E1/E2 prote
42	39	42.4	341	20 Y37104	Protein involved i
43	39	42.4	932	19 W53281	Murine phospholipa
44	39	42.4	933	20 W89200	Human phosphatidyl
45	39	42.4	933	21 Y44715	Human Phosphatidyl

ALIGNMENTS

RESULT 1

W63623
ID W63623 standard; Protein: 170 AA.

XX
AC W63623;

XX
DT 26-OCT-1998 (first entry)

XX
DE Human tumour necrosis factor receptor-6 beta protein.

XX
DE Human tumour necrosis factor receptor-6 alpha; TNFR-6 alpha; TNFR-6 beta; endothelial cells; keratinocytes; normal prostate; apoptosis; prostate tumour tissue.

XX
OS Homo sapiens.

XX
FH Key Location/Qualifiers

FT Peptide 1..30

FT Protein 31..170

FT Region note="TNFR-6 beta"

FT 31..166

FT /note= "Soluble extracellular domain"

XX
PN W09830694-A2.

XX
PD 16-JUL-1998.

XX
PF 13-JAN-1998; 98WO-US00153.

XX
PR 14-JAN-1997; 97US-0035496.

XX
PA (HUMA-) HUMAN GENOME SCI INC.

XX
PI Ebner R, Feng P, Gentz RL, Ni J, Ruben SM, Yu G;

DR WPI: 1998-399142/34.
 DR N-PSDB; V39086.
 XX Human tumour necrosis factor receptors 6-alpha and 6-beta - used in
 PT the diagnosis of immune system-related disorder(s)
 PT
 XX
 PS Claim 20; Fig 2A-2B; 91pp; English.
 XX
 CC The present sequence represents the human tumour necrosis factor
 CC receptor-6 beta (TNFR-6 beta) protein. The invention also provides
 CC for the TNFR-6 alpha protein (W63622). TNFR-6 alpha and TNFR-6 beta are
 CC members of the tumour necrosis factor receptor (TNFR) family. TNFRs
 CC are expressed in endothelial cells, keratinocytes, normal prostate and
 CC prostate tumour tissue. For a number of disorders of these cells,
 CC particularly of the immune system, substantially altered (whether
 CC increased or decreased) levels of TNFR-6 alpha and/or TNFR-6 beta gene
 CC expression can be detected, therefore the TNFR-6 alpha and TNFR-6 beta
 CC polypeptides, nucleic acids and antibodies are claimed to be useful in
 CC the diagnosis of such disorders. Mutations of the TNFR-6 alpha and
 CC TNFR-6 beta genes can also be detected. The TNFR polypeptides are
 CC also claimed to be useful for identifying ligands which may be useful
 CC in the treatment of apoptosis related disorders.
 XX
 SQ Sequence 170 AA;
 Query Match 100.0%; Score 92; DB 19; Length 170;
 Best Local Similarity 100.0%; Pred. NO. 2.4e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AETPTYPWRDAETGER 16
 Db | | | | | | | | | | | | | | | |
 31 aetptypwrdaetger 46
 RESULT 2
 Y28449
 ID Y28449 standard; Protein: 245 AA.
 XX
 AC Y28449;
 XX
 XX 29-SEP-1999 (first entry)
 DT
 DE A human tumour necrosis factor-R2-like proteins (TR2P)-1.
 XX
 KW Human tumour necrosis factor-R2-like protein: TR2P; achondroplasia;
 KW osteoporosis; developmental disorder; Cushing's syndrome;
 KW Charcot-Marie-Tooth disease; neurofibromatosis; hypothyroidism;
 KW hydrocephalus; seizure disorder; cerebellar palsy; spinal bifida;
 KW congenital glaucoma; cataract; sensorineural hearing loss;
 KW reproductive disorder; infertility; ovulatory defect; endometriosis;
 KW autoimmune disorder; ectopic pregnancy; teratogenesis; spermatogenesis;
 KW immunological disorder; AIDS; Addison's disease; allergy; bronchitis;
 KW atherosclerosis; diabetes mellitus; Chron's disease; lupus;
 KW irritable bowel syndrome; multiple sclerosis; infection;
 KW neoplastic disorder; adenocarcinoma; leukaemia; lymphoma; melanoma;
 KW myeloma; sarcoma.
 XX
 OS Homo sapiens.
 XX
 XX W09931128-A2.
 XX
 XX 24-JUN-1999.
 XX
 XX 02-DEC-1998; 98WO-US25649.
 XX
 XX 16-DEC-1997; 97US-0991945.
 XX
 XX (INCY-) INCYTE PHARM INC.
 XX
 XX Au-Young J, Bandman O, Hillman JL, Kaser MR, Tang YT;
 XX

DR WPI: 1999-457916/38.
 DR N-PSDB; X89503.
 XX New tumour necrosis factor-R2-like protein - useful in the treatment
 PT of osteogenesis, developmental, reproductive, immunological and
 PT neoplastic disorders
 PT
 XX
 PS Claim 1; Fig 1A-C; 81pp; English.
 XX
 CC The present sequence represents a human tumour necrosis factor-R2-like
 CC protein (TR2P). The protein is used to treat and prevent osteogenesis,
 CC developmental, reproductive, immunological and neoplastic disorders, and
 CC also to diagnose disorders associated with TR2 protein expression. Such
 CC disorders include osteogenesis disorders such as achondroplasia and
 CC osteoporosis, developmental disorders such as Cushing's syndrome,
 CC muscular dystrophy, and epilepsy, hereditary neuropathies such as
 CC Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism,
 CC hydrocephalus, seizure disorders such as cerebral palsy and spinal
 CC bifida, congenital glaucoma, cataract, or sensorineural hearing loss,
 CC reproductive disorders such as infertility, ovulatory defects and
 CC endometriosis, autoimmune disorders, ectopic pregnancy and teratogenesis,
 CC disruption of spermatogenesis, immunological disorders such as AIDS,
 CC Addison's disease, allergies, bronchitis, atherosclerosis, diabetes
 CC mellitus, Chron's disease, lupus and irritable bowel syndrome, multiple
 CC sclerosis, viral, fungal, helminthic, parasitic and protozoal infections,
 CC and neoplastic disorders including adenocarcinoma, leukaemia, lymphoma,
 CC melanoma, myeloma, sarcoma, and teratocarcinoma.
 XX
 SQ Sequence 245 AA;
 Query Match 100.0%; Score 92; DB 20; Length 245;
 Best Local Similarity 100.0%; Pred. NO. 3.5e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AETPTYPWRDAETGER 16
 Db | | | | | | | | | | | | | | | |
 31 aetptypwrdaetger 46
 RESULT 3
 Y42184
 ID Y42184 standard; Protein: 271 AA.
 XX
 AC Y42184;
 XX
 XX 17-DEC-1999 (first entry)
 DT
 DE Human mFLINT #1 protein sequence.
 XX
 KW Human: FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
 KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
 KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
 KW reperfusion-associated injury; aplastic anaemia; differentiation;
 KW growth; myelodysplastic syndrome; pancytopenic condition;
 KW myocardial ischaemia.
 XX
 OS Homo sapiens.
 XX
 XX W09950413-A2.
 XX
 XX 07-OCT-1999.
 XX
 XX 30-MAR-1999; 99WO-US06797.
 XX
 XX 30-MAR-1998; 98US-0079856.
 XX 20-MAY-1998; 98US-0086074.
 XX 09-SEP-1998; 98US-0099643.
 XX 17-DEC-1998; 98US-0112577.
 XX 18-DEC-1998; 98US-0112703.
 XX 18-DEC-1998; 98US-0112933.
 XX 22-DEC-1998; 98US-0113407.
 XX

PA (ELIL) LILLY & CO ELI.
 XX Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
 PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
 PI Song HY, Wang J, Wu X, Zuckerman SH;
 XX WPI; 1999-591319/50.
 DR N-PSDB; 225377.
 XX
 XX Use of mature FLINT for treating acute liver failure, inflammation,
 PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
 PT and proinflammatory activity
 XX
 XX Claim 31; Fig 3; 99pp; English.
 XX
 XX The present invention describes therapeutic applications of mature FLINT
 CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
 CC (mFLINT), which is a member of the tumour necrosis factor receptor
 CC superfamily, is used for treating acute liver failure, inflammation of
 CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
 CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
 CC injury or disorder such as hypercoagulation (including use with
 CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
 CC or disorder, Type I diabetes, cancer, cell damage or damage to an
 CC innocent bystander tissue that is induced by a chemotherapeutic agent or
 CC therapeutic irradiation, treating haematopoietic progenitor cells that
 CC have been exposed to therapeutic radiation or chemotherapy, aplastic
 CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
 CC also used for promoting the growth or differentiation of a haematopoietic
 CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
 CC resulting from abnormal myocardial ischaemia. The present sequence
 CC represents human mFLINT.
 XX
 XX Sequence 271 AA;

Query Match 100.0%; Score 92; DB 20; Length 271;
 Best Local Similarity 100.0%; Pred. No. 3.9e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AETPTYPWRDAETGER 16
 Db 2 aetptypwrdaetger 17
 |||||

RESULT 4
 Y42185
 ID Y42185 standard; Protein; 273 AA.
 AC Y42185;

DT 17-DEC-1999 (first entry)
 DE Human mFLINT #2 protein sequence.

XX Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
 KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
 KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
 KW reperfusion-associated injury; aplastic anaemia; differentiation;
 KW growth; myelodysplastic syndrome; pancytopenic condition;
 KW myocardial ischaemia.

OS Homo sapiens.
 XX WO9950413-A2.
 PN 07-OCT-1999.

PF 30-MAR-1999; 99WO-US06797.
 XX 30-MAR-1998; 98US-0079856.
 PR 20-MAY-1998; 98US-0086074.
 PR 09-SEP-1998; 98US-0099643.

PR 17-DEC-1998; 98US-0112577.
 PR 18-DEC-1998; 98US-0112703.
 PR 18-DEC-1998; 98US-0112933.
 PR 22-DEC-1998; 98US-0113407.
 XX (ELIL) LILLY & CO ELI.
 XX Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
 PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
 PI Song HY, Wang J, Wu X, Zuckerman SH;
 XX WPI; 1999-591319/50.
 DR N-PSDB; 225377.
 XX
 XX Use of mature FLINT for treating acute liver failure, inflammation,
 PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
 PT and proinflammatory activity
 XX
 XX Example 2; Fig 4; 99pp; English.
 XX
 XX The present invention describes therapeutic applications of mature FLINT
 CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
 CC (mFLINT), which is a member of the tumour necrosis factor receptor
 CC superfamily, is used for treating acute liver failure, inflammation of
 CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
 CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
 CC injury or disorder such as hypercoagulation (including use with
 CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
 CC or disorder, Type I diabetes, cancer, cell damage or damage to an
 CC innocent bystander tissue that is induced by a chemotherapeutic agent or
 CC therapeutic irradiation, treating haematopoietic progenitor cells that
 CC have been exposed to therapeutic radiation or chemotherapy, aplastic
 CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
 CC also used for promoting the growth or differentiation of a haematopoietic
 CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
 CC resulting from abnormal myocardial ischaemia. The present sequence
 CC represents human mFLINT.
 XX
 XX Sequence 273 AA;

Query Match 100.0%; Score 92; DB 20; Length 273;
 Best Local Similarity 100.0%; Pred. No. 3.9e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AETPTYPWRDAETGER 16
 Db 2 aetptypwrdaetger 17
 |||||

RESULT 5
 W66102
 ID W66102 standard; Protein; 300 AA.
 AC W66102;

DT 02-DEC-1998 (first entry)

DE Amino acid sequence of tumour necrosis related receptor (TR4).

XX Human; tumour necrosis related receptor; TR4; agonist; antagonist;
 KW inhibition; chronic; acute; inflammation; arthritis; septicemia;
 KW autoimmune disease; transplant rejection; stroke; cancer;
 KW Alzheimer's disease.

OS Homo sapiens.
 XX EP861850-A1.
 PN 02-SEP-1998.

PF 20-JAN-1998; 98EP-0300382.

Wed Jan 31 13:44:27 2001

```

PR 04-FEB-1997; 97US-0794796.
XX (SMIK ) SMITHKLINE BEECHAM CORP.
PA
XX Emery J, Tan KB, Truneh A, Young PR;
PI
XX WPI; 1998-508248/44.
DR N-PSDB; V07654.
XX
XX New DNA encoding tumour necrosis related receptor - used to treat
XX and prevent e.g. inflammation, arthritis, septicemia, autoimmune
XX diseases, transplant rejection, infection, stroke, ischaemia, AKDS,
XX restenosis, AIDS, bone disorders and cancer
XX
XX Claim 1; Fig 1; 21pp; English.
XX
XX This is the amino acid sequence of the human tumour necrosis related
XX receptor (TR4), used in the method of the invention. The TR4 protein
XX or its agonist can be used to treat a subject in need of enhanced
XX TR4 polypeptide activity. The antagonist is used to inhibit TR4
XX polypeptide activity. The active agents can be used for the
XX treatment and prevention of diseases such as chronic and acute
XX inflammation, arthritis, septicemia, autoimmune diseases, transplant
XX rejection, stroke, cancer, Alzheimer's disease.
XX
XX Sequence 300 AA;
XX
Query Match 100.0%; Score 92; DB 19; Length 300;
Best Local Similarity 100.0%; Pred. No. 4.3e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AETPTYPWRDAETGER 16
Db 31 aetptypwrdaetger 46
|||||
|||||

RESULT 6
ID W63622 standard; Protein; 300 AA.
XX
XX W63622;
XX
XX 26-OCT-1998 (first entry)
DT
XX
XX Human tumour necrosis factor receptor-6 alpha protein.
DE
XX
XX Human tumour necrosis factor receptor-6 alpha; TNFR-6 beta;
KW endothelial cells; keratinocytes; normal prostate; apoptosis;
KW prostate tumour tissue.
KW
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..30
XX Protein 31..300
XX /note= "TNFR-6 alpha"
XX Region 31..282
XX /note= "Soluble extracellular domain"
XX
XX W09830694-A2.
XX
XX 16-JUL-1998.
XX
XX 13-JAN-1998; 98WO-US00153.
XX
XX 14-JAN-1997; 97US-0035496.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ebner R, Feng P, Gentz RL, Ni J, Ruben SM, Yu G;
XX WPI; 1998-399142/34.
XX
DR

N-PSDB; V39085.
XX
XX Human tumour necrosis factor receptors 6-alpha and 6-beta - used in
XX the diagnosis of immune system-related disorder(s)
XX
XX Claim 20; Fig 1; 91pp; English.
XX
XX The present sequence represents the human tumour necrosis factor
XX receptor-6 alpha (TNFR-6 alpha) protein. The invention also provides
XX for the TNFR-6 beta protein (W63623). TNFR-6 alpha and TNFR-6 beta are
XX members of the tumour necrosis factor receptor (TNFR) family. TNFRs
XX are expressed in endothelial cells, keratinocytes, normal prostate and
XX prostate tumour tissue. For a number of disorders of these cells,
XX particularly of the immune system, substantially altered (whether
XX increased or decreased) levels of TNFR-6 alpha and/or TNFR-6 beta gene
XX expression can be detected, therefore the TNFR-6 alpha and TNFR-6 beta
XX polypeptides, nucleic acids and antibodies are claimed to be useful in
XX the diagnosis of such disorders. Mutations of the TNFR-6 alpha and
XX TNFR-6 beta genes can also be detected. The TNFR polypeptides are
XX also claimed to be useful for identifying ligands which may be useful
XX in the treatment of apoptosis related disorders.
XX
XX Sequence 300 AA;
XX
Query Match 100.0%; Score 92; DB 19; Length 300;
Best Local Similarity 100.0%; Pred. No. 4.3e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AETPTYPWRDAETGER 16
Db 31 aetptypwrdaetger 46
|||||
|||||

RESULT 7
ID Y03099 standard; Protein; 300 AA.
XX
XX Y03099;
XX
XX 09-DEC-1999 (first entry)
DT
XX
XX Human lung TNF-receptor protein.
DE
XX
XX Tumour necrosis factor; TNF; TNF receptor; human; lung; gene therapy;
KW detection; immunoassay; diagnosis; disease; immune system; tumour;
KW osteogenic system; cardiovascular system; central nervous system; asthma;
KW peripheral nervous systems; transplant incompatibility; antitumor;
KW rheumatoid arthritis; antiasthmatic; antiarthritic.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 134..1036
XX /*tag= a
XX /product= "TNF-receptor"
XX
XX DE19809978-A1.
XX
XX 16-SEP-1999.
XX
XX 09-MAR-1998; 98DE-1009978.
XX
XX 09-MAR-1998; 98DE-1009978.
XX
XX (BADI ) BASF AG.
XX
XX Kroeger B;
XX
XX WPI; 1999-519473/44.
XX
XX N-PSDB; Z09998.
XX
XX New soluble member of tumor necrosis factor receptor family, useful for
XX
XX

```


XX New tumour necrosis factor family receptor OPG-2

PS Claim 1: Page 18; 22pp; English.

XX The present sequence represents a mammalian tumour necrosis factor
CC receptor, designated OPG-2. OPG-2, is a member of the tumour necrosis
CC factor receptor family, and can be used: (i) to raise specific
CC antibodies (Ab), (ii) to treat osteopenic disorders associated with
CC excessive osteoclast activity, e.g. primary osteoporosis, Paget's
CC disease, hyperglycaemia of malignancy, or osteolytic metastases; (iii)
CC for affinity purification of cognate ligands, and (iv) to screen for
CC ligands (antagonists or agonists). Ab, or other OPG-2 blocking agents
CC such as soluble forms of the protein, are used to prevent, or reduce
CC severity of, an immune response, and for treating cancer. They can also
CC be used in diagnostic assays. The nucleic acid sequence encoding OPG-2
CC can be used as a probe to isolate related sequences from other species.

XX Sequence 300 AA;

Query Match 100.0%; Score 92; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 4.3e-07; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 0;

QY 1 AETPTYPRDRAETGER 16
Db 31 aetptyprdrdaetger 46

RESULT 10
Y06817
ID Y06817 standard; Protein; 300 AA.

AC Y06817;
XX
XX 24-JUN-1999 (first entry)
DT Human Dcr3 polypeptide.

XX Dcr3 polypeptide; tumour necrosis factor receptor; TNFR; Fas ligand;
KW apoptosis; T cell mediated immune response; allergy; asthma; cancer;
KW rheumatoid arthritis; Crohn's disease; guest vs. host disease; human;
KW gene therapy.

XX Homo sapiens.
OS
XX WO9914330-A1.
PN

XX 25-MAR-1999.
PD
XX 18-SEP-1998; 98WO-US19661.
PF

XX 30-JUL-1998; 98US-0094640.
PR
XX 18-SEP-1997; 97US-0059288.
PR
XX (GETH) GENENTECH INC.
PA

XX Ashkenazi AJ, Botstein D, Dodge KH, Goddard A, Gurney AL;
PI Kim KJ, Lawrence DA, Pitti R, Roy MA, Tumas DB;
PI Wood WI;

XX WPI: 1999-244032/20.
DR
XX N-PSDB; X32744.
DR

XX Dcr3 polypeptide related to tumor necrosis factor receptor

XX Claim 5: Fig 1; 88pp; English.

XX This represents a human Dcr3 polypeptide, a homologue of tumour necrosis
CC factor receptor (TNFR) polypeptide. Host cells containing a vector
CC comprising the Dcr3 nucleic acid can be used for the recombinant
CC expression of the protein. Dcr3 binds to Fas ligand, so it (or its

CC chimeras) are useful for modulating apoptosis in mammalian cells, also
CC other Fas-ligand induced activities, particularly to inhibit T cell
CC mediated immune responses, e.g. in treatment of allergy, asthma,
CC rheumatoid arthritis, Crohn's disease, guest vs. host disease etc. Dcr3
CC may also be used to identify specific binding proteins, potential
CC inhibitors. Antibodies against Dcr3 are used to treat cancer,
CC specifically of the lung and colon, also in diagnosis and for affinity
CC purification of the protein. Detecting mutations in the gene for Dcr3 is
CC also used to diagnose cancer, or predisposition to it. Dcr3 nucleic acid
CC is useful as hybridization probe to detect genomic or related sequences;
CC for chromosome and gene mapping; as source of antisense sequences; for
CC expression of recombinant Dcr3 and to generate transgenic animals (for
CC development and screening of therapeutic agents), also for in vivo or
CC ex vivo gene therapy.

XX Sequence 300 AA;

Query Match 100.0%; Score 92; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 4.3e-07; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 0;

QY 1 AETPTYPRDRAETGER 16
Db 31 aetptyprdrdaetger 46

RESULT 11
W97749
ID W97749 standard; Protein; 300 AA.

AC W97749;

XX 21-MAY-1999 (first entry)
DT Human tumour necrosis factor receptor zTNFR-5.

XX zTNFR-5; tumour necrosis factor receptor; TNFR; human;
KW cell maturation; bone cell regulation.

XX Homo sapiens.

XX Key Location/Qualifiers
FH Peptide /note= "signal peptide"
FT Protein /note= "mature protein"

FT Domain /note= "extracellular domain"
FT Region /note= "cysteine-rich pseudo-repeat 1"
FT Region /note= "cysteine-rich pseudo-repeat 1"

FT Region /note= "cysteine-rich pseudo-repeat 1"
FT Region /note= "cysteine-rich pseudo-repeat 1"
FT Region /note= "cysteine-rich pseudo-repeat 1"

XX WO9904001-A1.

XX 28-JAN-1999.

XX 21-JUL-1998; 98WO-US15072.

XX 21-JUL-1997; 97US-0053203.

XX (ZYMO) ZYMOGENETICS INC.

XX Farrah TM;

XX WPI: 1999-132245/11.

XX N-PSDB; X07226.

PT Novel tumour necrosis factor receptor ZTNFR5 - useful for
 PT regulating maturation of TNF-ligand bearing cells
 XX
 PS Claim 1; Page 84-85; 109pp; English.

CC This polypeptide comprises a new, secreted tumour necrosis factor
 CC receptor (see W97749), designated ZTNFR-5. Novel ZTNFR-5 encoding
 CC polynucleotides and polypeptides were initially identified by
 CC querying an expressed sequence tag (EST) database for sequences
 CC homologous to conserved motifs within the TNF receptor family.
 CC Based on this search, a contig of 16 ESTs (see X07226) was
 CC constructed. ZTNFR-5 polypeptides comprise 4 cysteine-rich repeats
 CC (see also W97750-55) that are homologous to other TNF receptors, in
 CC particular the soluble, secreted TNF receptor osteoprotegerin.
 CC ZTNFR-5 polypeptide can be prepared by recombinant methods. The
 CC polypeptide, especially the extracellular domain, can be used to
 CC generate a soluble variant of ZTNFR-5. The polypeptides and
 CC nucleic acids can be used to screen for ligands, agonists and
 CC antagonists of ZTNFR-5. The polypeptides can be used in bone cell
 CC regulation and to regulate the maturation of TNF ligand-bearing
 CC cells such as T- or B-cells, lymphocytes, peripheral blood
 CC mononuclear cells, polymorphonuclear leukocytes, fibroblasts or
 CC haematopoietic cells.
 XX
 SQ Sequence 300 AA;

Query Match 100.0%; Score 92; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 4.3e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AETPTYPWRDAETGER 16
 Db 31 aetpypwrdaetger 46
 |||||

RESULT 12
 W95082
 ID W95082 standard; Protein; 300 AA.

XX W95082;
 AC
 XX
 DT 20-MAY-1999 (first entry)
 XX
 XX Orphan receptor (HUMAN NTR-1) polypeptide.

XX HUMAN NTR-1; orphan receptor; osteoprotegerin; OPG; TNFR; human;
 KW tumour necrosis factor receptor; muscle disorder; bone mass; screening;
 KW muscle metabolism; binding agent; cognate ligand.

XX Homo sapiens.
 OS
 XX W09907738-A2.

PN 18-FEB-1999.

XX
 PD
 XX 04-AUG-1998; 98WO-US16202.

XX 06-AUG-1997; 97US-0054869.

XX (PROC) PROCTER & GAMBLE CO.
 PA (REG-) REGENERON PHARM INC.

XX Maslowski PJ, Morris J, Valenzuela DM;

XX WPI; 1999-167365/14.
 DR N-PSDB; X22300.

XX Novel orphan human receptor polypeptide and nucleic acid - useful as
 PT diagnostic reagents and for treatment of muscle disorders
 XX
 PS Claim 7; Page 21; 23pp; English.

CC This represents a HUMAN NTR-1 polypeptide, a novel orphan receptor. The
 CC protein is related to osteoprotegerin (OPG) and to tumour necrosis factor
 CC receptor (TNFR). Host cells transformed with a vector comprising the
 CC HUMAN NTR-1 nucleic acid are used for the recombinant expression of the
 CC protein. HUMAN NTR-1 proteins and antibodies immuno specific for the
 CC protein are useful for diagnosis and treatment of humans and animals,
 CC especially muscle disorders, as the receptor is involved in regulation of
 CC bone mass and muscle metabolism. HUMAN NTR-1 receptors are also useful
 CC for screening for novel binding agents, and cognate ligands, which may be
 CC used to treat disorders associated with HUMAN NTR-1 imbalance.

XX
 SQ Sequence 300 AA;

Query Match 100.0%; Score 92; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 4.3e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AETPTYPWRDAETGER 16
 Db 31 aetpypwrdaetger 46
 |||||

RESULT 13
 Y77458
 ID Y77458 standard; Protein; 300 AA.

XX Y77458;

XX 05-JUN-2000 (first entry)

XX Human TNF receptor-like protein, HDTEA84.

XX TNF receptor family; tumour necrosis factor; HDTEA84; HSLJD37R;
 KW Rank-like protein; RANKL; immune disorder; inflammation; allergy;
 KW immunosuppressant; antiarthritic; antirheumatoid; antiinflammatory;
 KW dermatological; antithyroid.

XX Homo sapiens.

XX W0200001817-A2.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-US12366.

XX 06-JUL-1998; 98US-0110938.

XX 13-JUL-1998; 98US-0114466.

XX 23-JUL-1998; 98US-0093897.

XX 12-AUG-1998; 98US-0132968.

XX 18-AUG-1998; 98US-0136214.

XX 11-SEP-1998; 98US-0099999.

XX (SCHE) SCHERING CORP.

XX WPI; 2000-171015/15.

XX N-PSDB; 292404.

XX New isolated mammalian genes, used to develop products for treating
 PT e.g. immune, inflammatory or allergic abnormalities, cancers or
 PT degenerative conditions -

XX Claim 24; Page 157; 218pp; English.

XX The invention relates to a number of primate and/or rodent proteins, and
 CC the genes which encode them. The invention encompasses human dendritic
 CC cell prostaglandin transporter (DC-PGT); the TNF (tumour necrosis
 CC factor) receptor family-related proteins HDTEA84, HSLJD37R and RANKL;
 CC human CC chemokine HCC5; human deubiquitinating proteins Dub11 and Dub

Wed Jan 31 13:44:27 2001

CC 12: human MD-1 and human and murine MD-2 proteins, which exhibit the
CC properties of ligands for proteins comprising a leucine-rich motif (LRR);
CC human cyclin E2; cDNAs encoding these proteins; and antibodies against
CC these proteins. The proteins can be used for modulating the physiology or
CC development of a cell. They can be used to mediate uptake of substrates
CC (e.g., prostaglandin-like molecules), to modulate or mediate cellular
CC interactions (e.g., induce or prevent trafficking, proliferation, or
CC differentiation of cells), or are intracellular proteins which are
CC important in various cellular processes such as the deubiquitination of
CC proteins or cell cycle regulation. The products can be used for
CC treating medical conditions such as immune, inflammatory or allergic
CC disorders, or abnormal cellular proliferation, for example, cancers or
CC degenerative conditions. They can be used to modulate immune responses in
CC disease states e.g., lupus erythematosus, Hashimoto's autoimmune
CC arthritis, systemic as well as acute and chronic inflammatory responses in which
CC thyroiditis, as well as expansion, and/or immunological T cell memory play an
CC important role. Sequences Y77458-Y77461 and Y77465-Y77468 represent
CC TNF receptor family-related proteins. Y77458 is the human protein
CC HDPEA84, Y77459-Y77461 are human HSLJD37R proteins, Y77465 is
CC murine Rank-like protein RANKL, and Y77466-Y77468 are human RANKL
CC proteins.
XX
SQ Sequence 300 AA;

Query Match 100.0%; Score 92; DB 21; Length 300;
Best Local Similarity 100.0%; Pred. No. 4.3e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AETPTYPWRDAETGER 16
Db 31 aetptypwrdaetger 46
|||||

RESULT 14
Y42183
ID Y42183 standard; Protein: 302 AA.
XX
AC Y42183;
XX
DT 17-DEC-1999 (first entry)

Human FLINT #2 protein sequence.

Human; FLINT; mFLINT; OPC3; tumour necrosis factor receptor; FasL;
KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
KW reperfusion-associated injury; aplastic anaemia; differentiation;
KW growth; myelodysplastic syndrome; pancytopenic condition;
KW myocardial ischaemia.

XX Homo sapiens.
XX WO9950413-A2.
XX 07-OCT-1999.
XX 30-MAR-1999; 99WO-US06797.
XX 30-MAR-1998; 98US-0079856.
XX 20-MAY-1998; 98US-0086074.
XX 09-SEP-1998; 98US-0099643.
XX 17-DEC-1998; 98US-0112577.
XX 18-DEC-1998; 98US-0112703.
XX 18-DEC-1998; 98US-0112933.
XX 22-DEC-1998; 98US-0113407.

XX (ELIL) LILLY & CO ELI.
XX Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
PI Hui KY, Kharitonkov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
PI Song HY, Wang J, Wu X, Zuckerman SH;

XX WPI; 1999-591319/50.
DR N-PSDB; Z25376.
XX use of mature FLINT for treating acute liver failure. Inflammation,
PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
PT and proinflammatory activity
XX
XX Example 2; Fig 2; 99pp; English.

XX The present invention describes therapeutic applications of mature FLINT
CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
CC (mFLINT), which is a member of the tumour necrosis factor receptor
CC superfamily, is used for treating acute liver failure, inflammation of
CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
CC injury or disorder such as hypercoagulation (including use with
CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
CC or disorder, Type I diabetes, cancer, cell damage or damage to an
CC innocent bystander tissue that is induced by a chemotherapeutic agent or
CC therapeutic irradiation, treating haematopoietic progenitor cells that
CC have been exposed to therapeutic radiation or chemotherapy, aplastic
CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
CC also used for promoting the growth or differentiation of a haematopoietic
CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
CC resulting from abnormal myocardial ischaemia. The present sequence
CC represents human FLINT.

XX Sequence 302 AA;

Query Match 100.0%; Score 92; DB 20; Length 302;
Best Local Similarity 100.0%; Pred. No. 4.3e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AETPTYPWRDAETGER 16
Db 31 aetptypwrdaetger 46
|||||

RESULT 15
W40382
ID W40382 standard; Protein: 355 AA.
XX
AC W40382;

XX 02-JUL-1998 (first entry)

XX S. glaucescens acbC protein.

XX Acarbose biosynthesis; acbA gene; acbB gene; acbC gene; acbD gene;
KW acbE gene; acbF gene; enzyme; alpha-amylase inhibitor; treatment;
KW diabetes.

XX Streptomyces glaucescens.

XX DEL9622783-AL.

XX 11-DEC-1997.

XX 07-JUN-1996; 96DE-1022783.

XX 07-JUN-1996; 96DE-1022783.

XX (FARH) HOECHST AG.

XX Decker H;

XX WPI; 1998-033827/04.

XX N-PSDB; T76903.

XX Recombinant DNA molecule comprising genes for biosynthesis of
PT acarbose - an alpha-amylase inhibitor useful in treatment of

PT diabetes
 XX
 PS
 XX Claim 4; Page 14-28; 35pp; German.
 CC This sequence represents the acbC protein from Streptomyces glaucescens
 CC GLA.O. This protein and the proteins encoded by the acbA, acbB, acbD,
 CC acbE and acbF genes are involved in the acarbose biosynthesis pathway.
 CC These genes and encoded enzymes are useful for producing acarbose, which
 CC is an alpha -amylase inhibitor useful in the treatment of diabetes.
 XX
 SQ Sequence 355 AA;

Query Match 53.3%; Score 49; DB 19; Length 355;
 Best Local Similarity 60.0%; Pred. No. 3.2;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Oy 1 AETPTYPRDAPTGE 15
 ||| ||||| : |
 Db 211 aetttrpwrtdgsae 225

Search completed: January 30, 2001, 16:45:21
 Job time: 581 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model
Run on: January 30, 2001, 16:50:16 ; Search time 149.64 Seconds
(without alignments)
7.260 Million cell updates/sec

Title: US-09-518-931-2_copy_31_46
Perfect score: 92
Sequence: 1 AETPTYPWRDAETGER 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues
Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_66: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45.5	49.5	5107	2 T29144	partial CDS - Caen
2	45	48.9	307	2 T17104	translation initia
3	44	47.8	168	2 T40352	single-stranded DN
4	43	46.7	922	2 T18878	hypothetical prote
5	42	45.7	256	2 H83565	conserved hypotet
6	42	45.7	1088	2 C83054	hypothetical prote
7	42	45.7	1239	1 Q0BE10	BOLFI protein - hu
8	42	45.7	1447	2 S50918	DNA helicase TP51
9	41.5	45.1	396	2 T15465	hypothetical prote
10	41	44.6	98	2 H75535	hypothetical prote
11	41	44.6	246	2 T04765	hypothetical prote
12	41	44.6	397	1 XNECY	aromatic-amino-aci
13	41	44.6	447	2 F69399	glutamate CoA-tr
14	41	44.6	534	2 T32020	hypothetical prote
15	41	44.6	557	2 T47665	beta-N-acetylhexos
16	41	44.6	1753	2 S30855	hypothetical prote
17	40	43.5	196	2 D70082	DNA-3-methyladenin
18	40	43.5	397	2 S71928	aromatic-amino-aci
19	40	43.5	439	2 S75545	hypothetical prote
20	40	43.5	532	2 A30766	beta-N-acetylhexos
21	40	43.5	609	2 A71461	hypothetical prote
22	40	43.5	659	2 T02838	probable membrane
23	40	43.5	683	2 A57037	dynein intermediat
24	40	43.5	832	2 T31792	hypothetical prote
25	40	43.5	850	2 S20482	RNA12 protein - ye
26	40	43.5	1350	2 G36793	hypothetical prote
27	40	43.5	2187	2 T30826	nascent polypeptid
28	40	43.5	3010	1 S18030	genome polyprotein
29	40	43.5	3562	2 A47171	chondroitin sulfat

30	39.5	42.9	654	2 A69656	methyl-accepting c
31	39.5	42.9	683	2 H72227	translation elonga
32	39	42.4	150	2 T00287	single-strand bind
33	39	42.4	163	2 T42184	single-stranded DN
34	39	42.4	166	2 T14869	single-stranded DN
35	39	42.4	174	2 A38487	single-stranded DN
36	39	42.4	174	2 B38487	single-stranded DN
37	39	42.4	174	2 S19955	single-stranded DN
38	39	42.4	175	1 DDEC1B	single-stranded DN
39	39	42.4	179	1 DDEC1F	single-stranded DN
40	39	42.4	186	2 H75492	probable N-6 adeni
41	39	42.4	282	2 E83283	hypothetical prote
42	39	42.4	321	2 A37842	hypothetical prote
43	39	42.4	327	2 T29031	hypothetical prote
44	39	42.4	365	2 JC4027	glutamate--ammonia
45	39	42.4	369	2 B71557	probable adenine 9

ALIGNMENTS

RESULT 1
T29144
partial CDS - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
C:Accession: T29144
R:Pauley, A.; Gattung, S.
submitted to the EMBL Data Library, July 1996
A:Description: The sequence of C. elegans cosmid K11C4.
A:Reference number: 220577
A:Accession: T29144
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-5107 <PAU>
A:Cross-references: EMBL:U64854; PIDN:AAB18318.1; GSPDB:GN00023; CESP:unc-68
A:Experimental source: strain Bristol N2; clone K11C4
C:Genetics:
A:Gene: CESP:unc-68
A:Map position: 5
A:Introns: 27/1; 64/3; 92/3; 127/1; 158/2; 1222/2; 1300/2; 1347/2; 1391/1; 1419/3; 15
/3; 3269/2; 3313/2; 3466/1; 3519/3; 3615/3; 3629/3; 3658/2; 3710/1; 3741/3; 3779/2; 3
C:Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 hom

Query Match 49.5%; Score 45.5; DB 2; Length 5107;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 10; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 2 ETPTYPWRDAETGE 15

DB 2414 ENPTPSKDAE-GE 2426

RESULT 2

T17104
translation initiation factor eIF-2 beta chain - apple tree (fragment)
C:Species: Malus domestica (apple tree)
C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 15-Sep-2000
C:Accession: T17104
R:Dong, Y.H.; Janssen, B.J.; Bielecki, L.L.; Atkinson, R.G.; Morris, B.A.; Gardner, R
J. Am. Soc. Hort. Sci. 122, 752-757, 1997
A:Title: Isolating and characterizing genes differentially expressed early in apple f
A:Reference number: Z18681
A:Accession: T17104
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-307 <DON>
A:Cross-references: EMBL:U80269; NID:gl732360; PID:gl732361
A:Experimental source: strain Granny Smith
C:Genetics:
A:Gene: eIF-2beta
C:Function:

Wed Jan 31 13:44:29 2001

A: Intron: 75/2; 98/1; 175/1; 276/2; 327/2; 457/3; 531/3; 635/1; 708/3; 721/3; 781/2;
C: Superfamily: Caenorhabditis elegans hypothetical protein C03C10.6

A: Description: catalyzes the exchange of eIF2-bound GDP for GTP
A: Pathway: protein biosynthesis
C: Superfamily: human translation initiation factor eIF-2 beta chain
C: Keywords: DNA binding; protein biosynthesis; zinc finger

Query Match 48.9%; Score 45; DB 2; Length 307;
Best Local Similarity 54.5%; Pred. No. 9.2;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 ETPTYPWRDAE 12
|||:||||:
Db 140 ETPYPWEGSD 150

RESULT 3
140352
single-stranded DNA-binding protein - Brucella abortus
N: Alternate names: helix-destabilizing protein
C: Species: Brucella abortus
C: Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 23-Jul-1999
C: Accession: I40352
C: Author: Y.; Oliveira, S.C.; Splitter, G.A.
I: Infect. Immun. 61, 5339-5344, 1993
A: Title: Isolation of Brucella abortus ssb and uvrA genes from a genomic library by use
A: Reference number: I40351; MUID: 94041662
A: Accession: I40352
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-168 <RES>
A: Cross-references: GB: L10843; NID: g144124; PIDN: AAA16818.1; PID: g144126
C: Genetics:
A: Gene: ssb
C: Superfamily: bacterial single-stranded DNA-binding protein; single-stranded DNA-binding
F: 21-100/Domain: single-stranded DNA-binding protein homology <SSD>

Query Match 47.8%; Score 44; DB 2; Length 168;
Best Local Similarity 77.8%; Pred. No. 7.1;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 WRDAETGER 16
|||:||||:
Db 40 WRDQTGER 48

RESULT 4
T18878
hypothetical protein C03C10.6 - Caenorhabditis elegans
C: Species: Caenorhabditis elegans
C: Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C: Accession: T18878; T20433
R: Berks, M.
submitted to the EMBL Data Library, August 1994
A: Reference number: Z19036
A: Accession: T18878
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-922 <WIL>
A: Cross-references: EMBL: Z35637; PIDN: CAA84690.1; GSPDB: GN00021; CESP: C03C10.6
A: Experimental source: clone C03C10
R: Gardner, A.
submitted to the EMBL Data Library, October 1994
A: Reference number: Z19274
A: Accession: T20433
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: DNA
A: Residues: 1-922 <WIL>
A: Cross-references: EMBL: Z38112; PIDN: CAA86235.1; GSPDB: GN00021; CESP: C03C10.6
A: Experimental source: clone E03A3
C: Genetics:
A: Gene: CESP: C03C10.6
A: Map position: 3

Query Match 46.7%; Score 43; DB 2; Length 922;
Best Local Similarity 54.5%; Pred. No. 60;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ETPTYPWRDAE 12
:|||||:
Db 330 DTPMPWHSTE 340

RESULT 5
H83565
conserved hypothetical protein PA0639 [imported] - Pseudomonas aeruginosa (strain PAO
C: Species: Pseudomonas aeruginosa
C: Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 15-Sep-2000
C: Accession: H83565
R: Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A: Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A: Reference number: A82950
A: Accession: H83565
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-256 <STO>
A: Cross-references: GB: AE004499; GB: AE004091; NID: g9946508; PIDN: AAG04028.1; GSPDB: GN
A: Experimental source: strain PA01
C: Genetics:
A: Gene: PA0639

Query Match 45.7%; Score 42; DB 2; Length 256;
Best Local Similarity 62.5%; Pred. No. 23;
Matches 10; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

QY 2 ETPTYPWRDA--ETGE 15
-||| || ||||
Db 144 ELPDYPDRDQWETGE 159

RESULT 6
C83054
hypothetical protein PA4735 [imported] - Pseudomonas aeruginosa (strain PA01)
C: Species: Pseudomonas aeruginosa
C: Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 15-Sep-2000
C: Accession: C83054
R: Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A: Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A: Reference number: A82950
A: Accession: C83054
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-1088 <STO>
A: Cross-references: GB: AE004487; GB: AE004091; NID: g9950991; PIDN: AAG08121.1; GSPDB: GN
A: Experimental source: strain PA01
C: Genetics:
A: Gene: PA4735

Query Match 45.7%; Score 42; DB 2; Length 1088;
Best Local Similarity 56.2%; Pred. No. 1e-02;
Matches 9; Conservative 1; Mismatches 2; Indels 4; Gaps 1;

QY 1 AETPTYPW---RDAAE 12
||||| |||:
Db 469 AETPAKPWNIVLRDAQ 484

```

RESULT 7
QB0E10
BOLFI protein - human herpesvirus 4 (strain B95-8)
C:Species: human herpesvirus 4, Epstein-Barr virus
C>Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 16-Jul-1999
C:Accession: A03041; A03752; S32995
R:Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrell, B.G.
Mol Biol Med. 1, 21-45, 1983
A:Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr virus
A:Reference number: A93065; MUID:85035713
A:Accession: A43041
A:Molecule type: DNA
A:Residues: 1-1239 <BAN>
A:Cross-references: EMBL:V01555; NID:g59074; PID:CAA24841.1; PID:gl334855
R:Baer, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.; H
Nature 310, 207-211, 1984
A:Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.
A:Reference number: A03794; MUID:84270667
A:Contents: annotation: protein coding region
C:Superfamily: human herpesvirus 4 BOLFI protein

Query Match 45.7%; Score 42; DB 1; Length 1239;
Best Local Similarity 53.3%; Pred. No. 1.2e+02;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AETPTYPWRDAETGE 15
| | | | | | | |
Db 217 AETNRYPWAAAGCGQ 231

RESULT 8
S50918
DNA helicase TPS1 - yeast (Saccharomyces cerevisiae)
N:Alternate names: chromosome segregation protein Sgs1; protein YM9646.02c; protein YMR1
C:Species: Saccharomyces cerevisiae
C>Date: 10-Feb-1995 #sequence_revision 12-May-1995 #text_change 22-Jun-1999
C:Accession: S50918; S51851; S55722; S55741; A56518; A56359
R:Pearson, D.; Bowman, S.
submitted to the EMBL Data Library, January 1995
A:Reference number: S50917
A:Accession: S50918
A:Molecule type: DNA
A:Residues: 1-1447 <PEA>
A:Cross-references: EMBL:Z47815; NID:g642280; PIDN:CAA87811.1; PID:g642282; MIPS:YM9190
R:Romeo, A.M.; Kleff, S.; Sternglanz, R.
submitted to the EMBL Data Library, December 1992
A:Reference number: S51851
A:Accession: S51851
A:Molecule type: DNA
A:Residues: 1-1447 <ROM>
A:Cross-references: EMBL:L07870; NID:g349193; PIDN:AAA35167.1; PID:g349194
R:Gangloff, S.; McDonald, J.P.; Bendixen, C.; Arthur, L.; Rothstein, R.
Mol. Cell. Biol. 14, 8391-8398, 1994
A:Title: The yeast type I topoisomerase Top3 interacts with Sgs1, a DNA helicase homolog
A:Reference number: A56359; MUID:95059068
A:Accession: S55722
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 679-998 <GAN>
A:Cross-references: EMBL:U22341
R:Gangloff, S.; McDonald, J.P.; Bendixen, C.; Arthur, L.; Rothstein, R.
submitted to the EMBL Data Library, March 1995
A:Reference number: S55741
A:Accession: S55741
A:Molecule type: DNA
A:Residues: 1-1447 <GAN>
A:Cross-references: EMBL:U22341; NID:g726277; PIDN:AAB60289.1; PID:g726279
R:Watt, P.M.; Louis, E.J.; Borts, R.H.; Hickson, I.D.
Cell 81, 253-260, 1995
A:Title: Sgs1: a eukaryotic homolog of Escherichia coli RecQ that interacts with topoisod

A:Reference number: A56518; MUID:95254634
A:Accession: A56518
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A:Molecule type: DNA
A:Residues: 1-1400 <WAT>
C:Genetics:
A:Gene: SGD:SGS1; TPS1
A:Cross-references: SGD:S0004802; MIPS:YMR190C
A:Map position: 13R
C:Superfamily: Bloom's syndrome helicase; DEAD/H box helicase homology; recQ helicase
C:Keywords: ATP; P-loop
F:700-993/Domain: DEAD/H box helicase homology <DEAD>
F:700-707/Region: nucleotide-binding motif A (P-loop)
F:804-809/Region: nucleotide-binding motif B
F:808-811/Region: DEAH motif
F:1033-1072/Domain: recQ helicase homology <RHH>

Query Match 45.7%; Score 42; DB 2; Length 1447;
Best Local Similarity 85.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 PTYPWRD 10
| | | | |
Db 658 PTYPWSD 664

RESULT 9
TI5465
hypothetical protein C09B8.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 10-Dec-1999
C:Accession: TI5465
R:Stelljes, L.
submitted to the EMBL Data Library, June 1995
A:Description: The sequence of C. elegans cosmid C09B8.
A:Reference number: S61138
A:Accession: TI5465
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-396 <STE>
A:Cross-references: EMBL:U29612; NID:g868273; PID:g868274; PIDN:AAA68800.1; CESP:C09B
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:C09B8.1
A:Introns: 31/3; 91/3; 167/2; 188/3; 239/1; 289/2
C:Superfamily: human inositol-polyphosphate 5-phosphatase

Query Match 45.1%; Score 41.5; DB 2; Length 396;
Best Local Similarity 61.5%; Pred. No. 43;
Matches 8; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Qy 4 PTYPW-RDAETGE 15
| | | | | |
Db 337 PTYPWSEDPENSE 349

RESULT 10
H75535
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: H75535
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896
A:Accession: H75535
A:Status: preliminary
A:Molecule type: DNA
```

Wed Jan 31 13:44:29 2001

```

J. Bacteriol. 169, 4710-4715, 1987
A:Title: Molecular analysis of the regulatory region of the Escherichia coli K-12 tyr
A:Cross-references: GB:AE001891; GB:AE000513; NID:G6457973; PIDN:AAF09888.1; PID:G645798
A:Reference number: I54888; MUID:86007418
A:Accession: I54888
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-39 <RE2>
A:Cross-references: GB:M17809; NID:G148086; PIDN:AAA24704.1; PID:G551844
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: E65213
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-397 <BLAT>
A:Cross-references: GB:U00096; NID:G2367339; PIDN:AAC77024.1; PID:G17904
A:Experimental source: strain K-12, substrain MG1655
A:Comment: This pyridoxal phosphate enzyme catalyzes the transamination reaction of t
C:Genetics:
A:Map position: 92 min
A:Gene: tyrB
A:Superfamily: aspartate aminotransferase
C:Keywords: aminotransferase; phosphoprotein; pyridoxal phosphate
F:247/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted
C:Genetics:
Query Match 44.6%; Score 41; DB 1; Length 397;
Best Local Similarity 58.3%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 ETPTYPRDAETGE 13
| | | | |
Db 144 EVSTYPMYDEAT 155

RESULT 13
F69399
glutamate CoA-transferase, subunit A (gctA) homolog - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: F69399
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes,
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch
A:Reference number: A69250; MUID:98049343
A:Accession: F69399
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-447 <KLE>
A:Cross-references: GB:AE001021; GB:AE000782; NID:G2689344; PIDN:AAB50043.1; PID:G264

Query Match 44.6%; Score 41; DB 2; Length 447;
Best Local Similarity 75.0%; Pred. No. 59;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ETPTYPR 9
| | | | |
Db 440 ERPVYPR 447

RESULT 14
T32020
hypothetical protein K02F6.9 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

A:Residues: 1-98 <NHI>
A:Cross-references: GB:AE001891; GB:AE000513; NID:G6457973; PIDN:AAF09888.1; PID:G645798
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR0301
A:Map position: 1

Query Match 44.6%; Score 41; DB 2; Length 98;
Best Local Similarity 58.3%; Pred. No. 12;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 4 PTYPRDAETGE 15
| | | | |
Db 32 PTYLYKDLNTGE 43

RESULT 11
T04765
hypothetical protein T16H5.200 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
C:Accession: T04765
R:Bevan, M.; De Haan, M.; Maarse, A.C.; Grivell, L.A.; Bancroft, I.; Mewes, H.W.; Mayer,
submitted to the Protein Sequence Database, June 1998
A:Reference number: T15383
A:Accession: T04765
A:Molecule type: DNA
A:Residues: 1-246 <BEV>
A:Cross-references: EMBL:AL024486
A:Experimental source: cultivar Columbia; BAC clone T16H5
C:Genetics:
A:Map position: 4
A:Introns: 82/3; 123/2
A:Note: T16H5.200

Query Match 44.6%; Score 41; DB 2; Length 246;
Best Local Similarity 50.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 ETPTYPRDAETGE 15
| | | | |
Db 195 ELPRYKWDVRVGE 208

RESULT 12
XNECY
aromatic-amino-acid transaminase (EC 2.6.1.57) tyrB [validated] - Escherichia coli
N:Alternate names: tyrosine aminotransferase
N:Contains: tyrosine transaminase (EC 2.6.1.5)
C:Species: Escherichia coli
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jun-2000
C:Accession: A30379; I52211; I54888; E65213; Q00614
R:Fotheringham, I.G.; Dacey, S.A.; Taylor, P.P.; Smith, T.J.; Hunter, M.G.; Finlay, M.E.
Biochem. J. 234, 593-604, 1986
A:Title: The cloning and sequence analysis of the aspC and tyrB genes from Escherichia
f E. coli with those of the pig aspartate aminotransferase isoenzymes.
A:Reference number: A30379; MUID:86242111
A:Accession: A30379
A:Molecule type: DNA
A:Residues: 1-397 <FOT>
A:Cross-references: EMBL:X03629
R:Kuramitsu, S.; Inoue, K.; Ogawa, T.; Ogawa, H.; Kagamiyama, H.
Biochem. Biophys. Res. Commun. 133, 134-139, 1985
A:Title: Aromatic amino acid aminotransferase of Escherichia coli nucleotide sequence of
A:Reference number: I52211; MUID:86076977
A:Accession: I52211
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-357 <RES>
A:Cross-references: GB:M12047; NID:G148084; PIDN:AAA24703.1; PID:G148085
R:Yang, J.; Pittard, J.

```

C:Accession: T32020
 R:Ledwith, J.; Wohldmann, P.; Rohlfing, T.
 submitted to the EMBL Data Library, July 1997
 A:Description: The sequence of C. elegans cosmid K02F6.
 A:Reference number: Z2112
 A:Accession: T32020
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-534 <LED>
 A:Cross-references: EMBL:AF016670; PIDN:AAB66107.1; GSPDB:GN00020; CESP:K02F6.9
 A:Experimental source: Strain Bristol N2; clone K02F6
 C:Genetics:
 A:Gene: CESP:K02F6.9
 A:Map position: 2
 A:Introns: 43/3; 77/3; 148/3; 200/3; 338/1; 424/3; 476/3

Query Match 44.6%; Score 41; DB 2; Length 534;
 Best Local Similarity 75.0%; Pred. No. 71;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 5 TYPWRDAE 12
 |||||
 Db 345 TYPWMDAD 352

RESULT 15
 T47665
 beta-N-acetylhexosaminidase-like protein - Arabidopsis thaliana
 N:Alternate names: protein T26112.140
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 19-May-2000
 C:Accession: T47665
 R:Monfort, A.; Casacuberta, E.; Puigdomenech, P.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.
 submitted to the Protein Sequence Database, February 2000
 A:Reference number: Z24471
 A:Accession: T47665
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-557 <MON>
 A:Cross-references: EMBL:AL132954
 A:Experimental source: cultivar Columbia; BAC clone T26112
 C:Genetics:
 A:Map position: 3
 A:Introns: 115/3; 143/3; 155/3; 193/1; 216/3; 260/2; 298/2; 332/1; 352/1; 363/2;
 A:Note: T26112.140
 C:Superfamily: beta-hexosaminidase

Query Match 44.6%; Score 41; DB 2; Length 557;
 Best Local Similarity 66.7%; Pred. No. 74;
 Matches 8; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

Oy 2 ETPTYYP--WRDA 11
 ||||| |:
 Db 232 ETPTYPNLWKG 243

Search completed: January 30, 2001, 16:50:19
 Job time: 709 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model
Run on: January 30, 2001, 17:05:52 ; Search time 80.04 seconds
(without alignments)
6.456 Million cell updates/sec

Title: US-09-518-931-2_COPY_31_46
Perfect score: 92
Sequence: 1 AETPTYPWRDAETGER 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	45	48.9	307	1 IF2B_MALDO	P55871 malus domes
2	44	47.8	168	1 SSB_BRUAB	Q07432 bruceella ab
3	43	46.7	922	1 YKL6_CAEEL	P42173 caenorhabdi
4	42	45.7	1239	1 V120_EBV	P03189 epstein-bar
5	42	45.7	1447	1 SGI1_YEAST	P35187 saccharomyc
6	41.5	45.1	396	1 I5P1_CAEEL	Q17848 caenorhabdi
7	41	44.6	397	1 TYRB_ECOLI	P04693 escherichia
8	41	44.6	447	1 YB99_ARCFU	O29069 archaeoglob
9	41	44.6	1753	1 YEW2_YEAST	P32634 saccharomyc
10	40	43.5	196	1 3MGH_BACSU	P94378 bacillus su
11	40	43.5	372	1 FWT_BOVIN	O77480 bos taurus
12	40	43.5	397	1 TYRE_SALTY	P74861 salmonella
13	40	43.5	420	1 HIS2_SYNP7	Q55267 synecococc
14	40	43.5	532	1 HEXA_DICDI	P13723 dictyosteli
15	40	43.5	601	1 Y858_CHLTR	O84866 chlamydia t
16	40	43.5	682	1 DV12_CHLRE	Q39578 chlamydomon
17	40	43.5	850	1 RN12_YEAST	P32843 saccharomyc
18	40	43.5	1350	1 W672_HSV11	Q00103 ictaluriid h
19	40	43.5	3562	1 PGCV_CHICK	Q90953 gallus gall
20	39.5	42.9	654	1 MCPB_BACSU	P54576 bacillus su
21	39.5	42.9	683	1 EFGL_THEMA	O9x1y4 thermotoga
22	39	42.4	90	1 SSB_SALTY	P37435 salmonella
23	39	42.4	173	1 SSB_PROMI	P28046 proteus mir
24	39	42.4	174	1 SSB7_ECOLI	P28044 escherichia
25	39	42.4	174	1 SSBP_ECOLI	P18022 escherichia
26	39	42.4	174	1 SSBP_ECOLI	P28045 escherichia
27	39	42.4	178	1 SSBF_ECOLI	P18310 escherichia
28	39	42.4	379	1 YX1L_ANASP	P29978 anabaena sp
29	39	42.4	615	1 MTHR_CAEEL	Q17693 caenorhabdi
30	39	42.4	615	1 NTNO_BOVIN	P51143 bos taurus
31	39	42.4	677	1 SACC_BACSU	P05656 bacillus su
32	39	42.4	741	1 PIFA_ECOLI	P96329 escherichia
33	39	42.4	891	1 POL2_BAMMN	P89684 barley mild

34	39	42.4	933	1	PLD2_RAT	P70498 rattus norv
35	38.5	41.8	311	1	IR01_HCMVA	P09710 human cytom
36	38.5	41.8	473	1	RBL1_THIFE	P28895 thiobacillu
37	38.5	41.8	606	1	ACEA_MYCLE	P46831 mycobacteri
38	38	41.3	95	1	HXDA_CHICK	P24341 gallus gall
39	38	41.3	172	1	RS4_ARCFU	O28000 archaeoglob
40	38	41.3	174	1	SSB_RHIME	P56898 rhizobium m
41	38	41.3	210	1	FTSQ_AGRU	O30990 agrobacteri
42	38	41.3	260	1	CY12_SOLTU	P29610 solanum tub
43	38	41.3	340	1	HXDA_HUMAN	P28358 homo sapien
44	38	41.3	472	1	DAX1_MOUSE	Q01066 mus musculu
45	38	41.3	604	1	HMDH_NICSY	Q01559 nicotiana s

ALIGNMENTS

```
RESULT 1
IF2B_MALDO
ID IF2B_MALDO STANDARD; PRT; 307 AA.
AC P55871;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 01-OCT-2000 (Rel. 40, Last annotation update)
DE EUKARYOTIC TRANSLATION INITIATION FACTOR 2 BETA SUBUNIT (EIF-2-BETA)
DE (FRAGMENT).
OS Malus domestica (Apple) (Malus sylvestris)
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurossids I;
OC Rosales; Rosaceae; Malus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. GRANNY SMITH;
RA Dong Y.H., Janssen B.J., Bielecki L.L., Atkinson R.G.,
RA Morris B.A., Gardner R.C.;
RA Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: EIF-2 FUNCTIONS IN THE EARLY STEPS OF PROTEIN SYNTHESIS
CC BY FORMING A TERNARY COMPLEX WITH GTP AND INITIATOR TRNA. THIS
CC COMPLEX BINDS TO A 40S RIBOSOMAL SUBUNIT, FOLLOWED BY MRNA BINDING
CC TO FORM A 43S PREINITIATION COMPLEX. JUNCTION OF THE 60S RIBOSOMAL
CC SUBUNIT TO FORM THE 80S INITIATION COMPLEX IS PRECEDED BY
CC HYDROLYSIS OF THE GTP BOUND TO EIF-2 AND RELEASE OF AN EIF-2-GDP
CC BINARY COMPLEX. IN ORDER FOR EIF-2 TO RECYCLE AND CATALYZE ANOTHER
CC ROUND OF INITIATION, THE GDP BOUND TO EIF-2 MUST EXCHANGE WITH GTP
CC BY WAY OF A REACTION CATALYZED BY EIF-2B (BY SIMILARITY).
CC -!- SUBUNIT: HETEROTRIMER COMPOSED OF AN ALPHA, A BETA, AND A GAMMA
CC CHAIN.
CC -!- SIMILARITY: BELONGS TO THE EIF2B / EIF5 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; U80269; AAC0384.1; -.
CC INTERPRO; IPR002735; -.
CC PFAM; PF01873; eIF5_eIF2B; 1.
CC Initiation factor; Protein biosynthesis; Zinc-finger.
CC NON_TER 1 1
CC DOMAIN 30 33 POLY-ARG (BASIC).
CC DOMAIN 60 65 POLY-LYS (BASIC).
CC DOMAIN 101 106 POLY-LYS (BASIC).
CC ZN_FING 259 283 C4-TYPE (POTENTIAL).
CC SEQUENCE 307 AA; 34620 MW; C2BFCEBBD1CEC210 CRC64;

```

Query Match 48.9%; Score 45; DB 1; Length 307;
Best Local Similarity 54.5%; Pred. No. 4.3;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Wed Jan 31 13:44:30 2001

```

RA  Gardner A., Berks M.;
RL  Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; Z38112; CAA86235.1;
DR  EMBL; Z35637; CAA86235.1; JOINED.
DR  EMBL; Z35637; CAA84690.1;
DR  EMBL; Z38112; CAA84690.1; JOINED.
DR  WORMPEP; C03C10.6; C0202959.
KW  Hypothetical protein.
SQ  SEQUENCE 922 AA; 103818 MW; DE0808AB97FC87D8 CRC64;

Query Match 46.7%; Score 43; DB 1; Length 922;
Best Local Similarity 54.5%; Pred. No. 27;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY  2 ETPYTPWRDAE 12
DB  330 DTPWYPMHSTE 340
      :||| |||
      :||| |||

RESULT 4
V120_EBV STANDARD; PRT; 1239 AA.
AC  P03189;
DT  21-JUL-1986 (Rel. 01, Created)
DT  21-JUL-1986 (Rel. 01, Last sequence update)
DT  01-OCT-1996 (Rel. 34, Last annotation update)
DE  CAPSID ASSEMBLY PROTEIN BOLF1.
GN  BOLF1.
OS  Epstein-barr virus (strain B95-8) (Human herpesvirus 4).
OC  Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC  Gammaherpesvirinae; Lymphocryptovirus.
RN  [1]
RP  MEDLINE=84270667; PubMed=6087149;
RX  Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
RA  Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
RT  Tuffnell P.S., Barrell B.G.;
RT  "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
RL  Nature 310:207-211(1984).
CC  -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL37,
CC  EBV-1 23, EBV BOLF1, VZV 21, HVS-1 63, AND HCMV UL47.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; V01555; CAA24841.1;
DR  PIR; A03752; Q0BE10.
DR  PIR; S32995; S32995.
KW  Capsid assembly.
SQ  SEQUENCE 1239 AA; 132748 MW; 6C5DBFC55F2FF729 CRC64;

Query Match 45.7%; Score 42; DB 1; Length 1239;
Best Local Similarity 53.3%; Pred. No. 53;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY  1 AETPTYPWRDAETGE 15
      ||| ||| |||

RA  Gardner A., Berks M.;
RL  Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; Z38112; CAA86235.1;
DR  EMBL; Z35637; CAA86235.1; JOINED.
DR  EMBL; Z35637; CAA84690.1;
DR  EMBL; Z38112; CAA84690.1; JOINED.
DR  WORMPEP; C03C10.6; C0202959.
KW  Hypothetical protein.
SQ  SEQUENCE 922 AA; 103818 MW; DE0808AB97FC87D8 CRC64;

Query Match 46.7%; Score 43; DB 1; Length 922;
Best Local Similarity 54.5%; Pred. No. 27;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY  2 ETPYTPWRDAE 12
DB  330 DTPWYPMHSTE 340
      :||| |||
      :||| |||

RESULT 3
YKL6_CAEEL STANDARD; PRT; 922 AA.
ID  YKL6_CAEEL
AC  P42173;
DT  01-NOV-1995 (Rel. 32, Created)
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DT  30-MAY-2000 (Rel. 39, Last annotation update)
DE  HYPOTHETICAL 103.8 KDA PROTEIN C03C10.6 IN CHROMOSOME III.
GN  C03C10.6.
OS  Caenorhabditis elegans.
OC  Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC  Rhabditidae; Peloderinae; Caenorhabditis.
RN  [1]
RP  SEQUENCE FROM N.A.
RT  STRAIN=BRISTOL N2;

Query Match 47.8%; Score 44; DB 1; Length 168;
Best Local Similarity 77.8%; Pred. No. 3.3;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY  8 WRDAETGER 16
DB  40 WRDROTGER 48
      ||| ||| |||

RESULT 3
YKL6_CAEEL STANDARD; PRT; 922 AA.
ID  YKL6_CAEEL
AC  P42173;
DT  01-NOV-1995 (Rel. 32, Created)
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DT  30-MAY-2000 (Rel. 39, Last annotation update)
DE  HYPOTHETICAL 103.8 KDA PROTEIN C03C10.6 IN CHROMOSOME III.
GN  C03C10.6.
OS  Caenorhabditis elegans.
OC  Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC  Rhabditidae; Peloderinae; Caenorhabditis.
RN  [1]
RP  SEQUENCE FROM N.A.
RT  STRAIN=BRISTOL N2;

```


Db 217 AETNRYPWAAAGQG 231

RESULT 5

```
SGSI_YEAST
ID SGSI_YEAST STANDARD; PRT; 1447 AA.
AC P35187;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE HELICASE SGSI (HELICASE TPS1).
GN SGSI OR TPS1 OR YMR190C OR YMR646.02C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC Saccharomycetaceae; Saccharomycetes.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W303;
RX MEDLINE=95059068; PubMed=7969174;
RA Gangloff S., McDonald J.P., Bendixen C., Arthur L., Rothstein R.;
RT "The yeast type I topoisomerase Top3 interacts with Sgs1, a DNA
RT helicase homolog: a potential eukaryotic reverse gyrase.";
RL Mol. Cell. Biol. 14:8391-8398(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Romeo A.M., Kleff S., Sternglanz R.;
RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Pearson D., Bowman S., Barrell B.G., Rajandream M.A.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INTERACTS WITH TOPOISOMERASES II AND TOP3. COULD CREATE
CC A DELETTERIOUS TOPOLOGICAL SUBSTRATE THAT TOP3 PREFERENTIALLY
CC RESOLVES. THE TOP3-SGSI PROTEIN COMPLEX MAY FUNCTION AS A
CC EUKARYOTIC REVERSE GYRASE INTRODUCING POSITIVE SUPERCOILS INTO
CC EXTRACHROMOSOMAL RIBOSOMAL DNA RINGS.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- SIMILARITY: BELONGS TO THE RECO SUBFAMILY OF HELICASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; L07870; AAA35167.1; -
EMBL; U22341; AAB60289.1; -
EMBL; Z47815; CAA87811.1; -
SGD; S0004802; SGSI.
DR INTERPRO; IPR001410; -
DR INTERPRO; IPR001650; -
DR INTERPRO; IPR002121; -
DR INTERPRO; IPR002464; -
DR PFAM; PF00270; DEAD; 1.
DR PFAM; PF00570; HRDC; 1.
DR PFAM; PF00271; helicase_C; 1.
DR PROSITE; PS00690; DEAH_ATP_HELICASE; UNKNOWN_1.
KW Helicase; ATP-binding; Nuclear protein.
FT DOMAIN 631 639 ASP/GLU-RICH (HIGHLY ACIDIC).
FT NP_BIND 714 721 ATP (BY SIMILARITY).
FT SITE 808 811 DEAH BOX.
SQ SEQUENCE 1447 AA; 163836 MW; 0DC320B41756A3C3 CRC64;
```

Query Match 45.7%; Score 42; DB 1; Length 1447;
Best Local Similarity 85.7%; Pred. No. 62;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 PTPWPD 10
|||||

Db 658 PTPWSD 664

RESULT 6

```
ISPI_CAEEL
ID ISPI_CAEEL STANDARD; PRT; 396 AA.
AC Q17848;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE PROBABLE TYPE I INOSITOL-1,4,5-TRISPHOSPHATE 5-PHOSPHATASE
DE (EC 3.1.3.56) (5PTASE).
GN C09B8.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Stellyes L.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: D-MYO-INOSITOL 1,4,5-TRISPHOSPHATE + H(2)O -
CC D-MYO-INOSITOL 1,4-BISPHOSPHATE + PHOSPHATE.
CC -!- SIMILARITY: BELONGS TO THE INOSITOL-1,4,5-TRISPHOSPHATE
CC 5-PHOSPHATASE TYPE I FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; U29612; AAA68800.1; -
DR WORMPEP; C09B8.1; CE02461.
DR INTERPRO; IPR000300; -
DR PFAM; PF00783; IPPC; 1.
DR Hydrolase.
KW SEQUENCE 396 AA; 46085 MW; 1FEC5982B00ACC5B CRC64;
```

Query Match 45.1%; Score 41.5; DB 1; Length 396;
Best Local Similarity 61.5%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Qy 4 PTPW-RDAETGE 15
|||||

RESULT 7

```
TYRB_ECOLI
ID TYRB_ECOLI STANDARD; PRT; 397 AA.
AC P04693;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE AROMATIC-AMINO-ACID AMINOTRANSFERASE (EC 2.6.1.57) (ARAT).
GN TYRB.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8624111; PubMed=3521591;
RA Fotheringham I.G., Dacey S.A., Taylor P.P., Smith T.J., Hunter M.G.,
RA Finlay M.E., Primrose S.B., Parker D.M., Edwards R.M.;
RT "The cloning and sequence analysis of the aspC and tyrB genes from
RT Escherichia coli K12. Comparison of the primary structures of the
RT aspartate aminotransferase and aromatic aminotransferase of E. coli
RT with those of the pig aspartate aminotransferase isoenzymes.";
RL Biochem. J. 234:593-604(1986).
```

	RESULT	9			
	YEW2_YEAST				
ID	YEW2_YEARST		STANDARD;	PRT; 1753 AA.	
AC	P32634;				
DT	01-OCT-1993	(Rel. 27,	Created)		
DT	01-OCT-1993	(Rel. 27,	Last sequence update)		
DT	01-FEB-1995	(Rel. 31,	Last annotation update)		

DE HYPOTHETICAL 195.4 KDA PROTEIN IN RPS26B-GLC7 INTERGENIC REGION.
GN YER132C OR SVGP-ORF50.
OS Saccharomyces cerevisiae (Baker's yeast).
CC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
CC Saccharomycetaceae; Saccharomycetes.
RN [1]
RP SEQUENCE FROM N.A.
RA Mulligan J.T., Dietrich F.S., Hennessey K.M., Allen E., Araujo R.,
RA Wei Y., Taylor P., Nakahara K., Roberts D., Davis R.W.;
RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
RA Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,
RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
RA Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P., Oh C.,
RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO S.POMBE RAL2.
CC -!- SIMILARITY: TO YEAST MDS3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U18916; AAC03230.1; -
DR PIR: S30855; S30855.
DR SGD: S0000934; PMD1.
KW Hypothetical protein.
SQ SEQUENCE 1753 AA; 195382 MW; E42529986C2508B CRC64;

Query Match 44.6%; Score 41; DB 1; Length 1753;
Best Local Similarity 50.0%; Pred. NO. 1.le=02;
Matches 10; Conservative 1; Mismatches 5; Indels 4; Gaps 1;

QY 1 AETPT----YPRDAETGER 16
DB 277 AEIPTLALYPQREANTSRR 296

RESULT 10
ID 3MGL_BACSU STANDARD; PRT; 196 AA.
AC P94378;
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PUTATIVE 3-METHYLAENINE DNA GLYCOSYLASE (EC 3.2.2.-).
GN YXLJ.
OS Bacillus subtilis.
CC Bacteria; Firmicutes; Bacillus/Clostridium group;
CC Bacillus/Staphylococcus group; Bacillus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / BGSC1A1;
RX MEDLINE=97124196; PubMed=8969509;
RA Yoshida K., Shindo K., Sano H., Seki S., Fujimura M., Yanai N.,
RA Miwa Y., Fujita Y.;
RT "Sequencing of a 65 kb region of the Bacillus subtilis genome
RT containing the lic and cel loci, and creation of a 177 kb contig
RT covering the gnt-sackY region.";
RL Microbiology 142:3113-3123(1996).
CC -!- SIMILARITY: BELONGS TO THE MPG FAMILY OF DNA GLYCOSYLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D83026; BAA11741.1; -
DR EMBL: Z99123; CAB15888.1; -
DR SUBTILIST; BG12555; YXLJ.
KW Hypothetical protein; DNA repair; Hydrolase.
SQ SEQUENCE 196 AA; 22124 MW; E63A02E6285E4FF0 CRC64;

Query Match 43.5%; Score 40; DB 1; Length 196;
Best Local Similarity 53.3%; Pred. NO. 17;
Matches 8; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 ETPTYPWRDAETGER 16
DB 178 EARDYPWRFWVTGNR 192

RESULT 11
ID FMT_BOVIN STANDARD; PRT; 372 AA.
AC O77480;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE METHIONYL-TRNA FORMYLTRANSFERASE, MITOCHONDRIAL PRECURSOR (EC 2.1.2.9)
DE (MTFMT) (FRAGMENT).
OS Bos taurus (Bovine).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
CC Bovidae; Bovinae; Bos.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 16-22; 68-77 AND 177-186.
RC TISSUE=HEART AND LIVER;
RX MEDLINE=98279026; PubMed=9614118;
RA Takeuchi N., Kawakami M., Omori A., Ueda T., Spremulli L.L.,
RA Watanabe K.;
RT "Mammalian mitochondrial methionyl-trna transformylase from bovine
RT liver. Purification, characterization, and gene structure.";
RL J. Biol. Chem. 273:15085-15090(1998).
CC -!- FUNCTION: FORMYLATES METHIONYL-TRNA IN MITOCHONDRIA. A SINGLE
CC TRNA(MET) GENE GIVES RISE TO BOTH AN INITIATOR AND AN ELONGATOR
CC SPECIES VIA AN UNKNOWN MECHANISM.
CC -!- CATALYTIC ACTIVITY: 10-FORMYLTETRAHYDROFOLATE + L-METHIONYL-TRNA
CC + H(2)O = TETRAHYDROFOLATE + N-FORMYLMETHIONYL-TRNA.
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC -!- DOMAIN: COMPOSED OF AN N- AND A C-TERMINAL DOMAIN. THE N-TERMINAL
CC DOMAIN CARRIES THE TETRAHYDROFOLATE (THF)-BINDING SITE AND THE C-
CC TERMINAL DOMAIN IS PRESUMABLY INVOLVED IN POSITIONING THE MET-TRNA
CC SUBSTRATE FOR THE FORMYLATION REACTION.
CC -!- SIMILARITY: BELONGS TO THE FMT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB004316; BAA31237.1; -
DR INTERPRO: IPR002376; -
DR PFAM: PF00551; formyl_transf; 1.
KW Transferrase; Methyltransferase; Protein biosynthesis; Mitochondrion;
KW Transit peptide.
FT NON_TER 1 1
FT TRANSIT <1 15
FT CHAIN 16 372
FT MITOCHONDRION (POTENTIAL).
FT METHIONYL-TRNA FORMYLTRANSFERASE.

Wed Jan 31 13:44:30 2001

us-09-518-931-2_copy_31_46.rsp

```

SQ SEQUENCE 372 AA: 41750 MW: BFB57979ADF8AC35 CRC64;

Query Match 43.5%; Score 40; DB 1; Length 372;
Best Local Similarity 40.0%; Pred. No. 32;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 AETPTYPWRDAETGE 15
   : | | | | |
DB 89 SLPVYEPDVGSGE 103

RESULT 13
HISZ_SYN7 STANDARD; PRT; 420 AA.
ID HISZ_SYN7
AC Q55267;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ATP PHOSPHORIBOSYLTRANSFERASE REGULATORY SUBUNIT.
GN HISZ.
OS Synecococcus sp. (strain PCC 7942) (Anacystis nidulans R2).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95050240; PubMed=7961432;
RT Tsinoremas N.F., Kutach A.K., Strayer C.A., Golden S.S.;
RT "Efficient gene transfer in Synechococcus sp. strains PCC 7942 and
RT PCC 6301 by interspecies conjugation and chromosomal recombination.";
RL J. Bacteriol. 176:6764-6768(1994).
CC -!- FUNCTION: MAY ALLOW THE REGULATION OF ATP
CC PHOSPHORIBOSYLTRANSFERASE ACTIVITY BY HISTIDINE (BY SIMILARITY).
CC -!- PATHWAY: FIRST STEP IN HISTIDINE BIOSYNTHETIC PATHWAY. IS VERY
CC IMPORTANT IN THE REGULATION OF HISTIDINE METABOLISM.
CC -!- SUBUNIT: BINDS TO HISG (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- MISCELLANEOUS: THIS FUNCTION IS GENERALLY FULFILLED BY THE C-
CC TERMINAL PART OF HISG, WHICH IS MISSING IN SOME BACTERIA SUCH AS
CC THIS ONE.
CC -!- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
CC STRONG, TO HISTIDYL-TRNA SYNTHETASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L35476; AAA64445.1; -
CC INTERPRO: IPR002106; -
CC PROSITE; PS00179; AA_TRNA_LIGASE_II_1; FALSE_NEG.
CC PROSITE; PS00339; AA_TRNA_LIGASE_II_2; FALSE_NEG.
CC Histidine biosynthesi.
KW SEQUENCE 420 AA; 46568 MW; C493F87E8777C961 CRC64;
SQ

Query Match 43.5%; Score 40; DB 1; Length 420;
Best Local Similarity 61.5%; Pred. No. 37;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 ETPTYPWRDAETG 14
   | | | | | | |
DB 284 ETPTGPWVLAQGG 296

RESULT 14
HEXA_DICDI STANDARD; PRT; 532 AA.
ID HEXA_DICDI
AC P13723;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE BETA-HEXOSAMINIDASE A PRECURSOR (EC 3.2.1.52) (N-ACETYL-BETA-
DE GLUCOSAMINIDASE) (BETA-N-ACETYLHEXOSAMINIDASE).
GN NAGA.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Dictyostelida; Dictyostelium.
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=89034175; PubMed=2972716;
RX Graham T.R., Zassenhaus H.P., Kaplan A.;
RT "Molecular cloning of the cDNA which encodes beta-N-
RT acetylhexosaminidase A from Dictyostelium discoideum. Complete amino

```

RT acid sequence and homology with the human enzyme.";
RL J. Biol. Chem. 263:16823-16829(1988).
CC -!- FUNCTION: BETA-HEXOSAMINIDASE A IS RESPONSIBLE FOR THE DEGRADATION
CC OF GM2 GANGLIOSIDES, AND A VARIETY OF OTHER MOLECULES CONTAINING
CC TERMINAL N-ACETYL HEXOSAMINES. THIS ENZYME PLAYS A ROLE DURING THE
CC SLUG STAGE OF DEVELOPMENT IN THE MAINTENANCE OF PSEUDOPUSMODIA OF
CC NORMAL SIZE.
CC -!- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL NON-REDUCING N-ACETYL-
CC D-HEXOSAMINE RESIDUES IN N-ACETYL-BETA-D-HEXOSAMINIDES.
CC -!- SUBCELLULAR LOCATION: LYSOSOMAL.
CC -!- SIMILARITY: BELONGS TO FAMILY 20 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: J04065; AAA33230.1; -
DR PIR: A30766; A30766.
DR PIR: A31778; A31778.
DR HSSP: P07686; IQBD.
DR DICTYDB: DD02013; NAGA.
DR INTERPRO: IPR001540; -
DR PFAM: PF00728; Glyco_hydro_20; 1.
DR PRINTS: PR00738; GLYDRLASE20.
KW Hydrolase; Glycosidase; Lysosome; Signal; Glycoprotein.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 532 BETA-HEXOSAMINIDASE A.
FT CARBOHYD 72 72 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 79 79 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 427 427 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT ACT_SITE 308 308 CATALYTIC ACID (BINDS TO THE GLYCOSIDIC
FT LINKAGE) (BY SIMILARITY).
SQ SEQUENCE 532 AA; 59787 MW; C5651DD79F456514 CRC64;

Query Match 43.5%; Score 40; DB 1; Length 532;
Best Local Similarity 55.6%; Pred. No. 46;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AETPTYPWR 9
Db 150 SDSPPYPWR 158
:::| | | | |

RESULT 15
Y858.CHLTR
ID Y858.CHLTR STANDARD; PRT; 601 AA.
AC 084866;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYPOTHETICAL PROTEIN CT858.
GN CT858.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/UW-3/CX;
RX MEDLINE=99000809; PubMed=9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis.";
RL Science 282:754-759(1998).
CC -!- SIMILARITY: TO C.PNEUMONIAE CPN1016.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE001359; AAC68456.2; -
DR Hypothetical protein.
KW SEQUENCE 601 AA; 67252 MW; 2D6B3068C546205B CRC64;
SQ

Query Match 43.5%; Score 40; DB 1; Length 601;
Best Local Similarity 66.7%; Pred. No. 53;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 PTPYWRDAE 12
Db 323 PTYSWQDME 331
||| | | | |

Search completed: January 30, 2001, 17:05:55
Job time: 1216 sec

Wed Jan 31 13:44:30 2001

us-09-518-931-2_copy_31_46.rsp

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:54:36 ; Search time 254.27 Seconds
(without alignments)
7.375 Million cell updates/sec

Title: US-09-518-931-2_COPY_31_46
Perfect score: 92
Sequence: 1 AETPTVPRDAETGER 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues
Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_15:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phage:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_undefined:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	92	100.0	300	4 095407	095407 homo sapien
2	49	53.3	537	3 042801	042801 aspergillus
3	45.5	49.5	5071	5 P91905	P91905 caenorhabdi
4	45.5	49.5	5107	5 094279	094279 caenorhabdi
5	45	48.9	310	11 Q9QY29	Q9QY29 mus musculu
6	45	48.9	7257	2 Q9L8C7	Q9L8C7 sorangium c
7	44.5	48.4	1521	5 Q9W0P8	Q9W0P8 drosophila
8	44	47.8	168	10 Q22052	Q22052 acetabulari
9	43.5	47.3	1421	2 Q9L8C9	Q9L8C9 sorangium c
10	43.5	47.3	1421	2 Q9KJ00	Q9KJ00 sorangium c
11	43	46.7	180	4 Q9NW10	Q9NW10 homo sapien
12	43	46.7	5532	2 Q9S0R4	Q9S0R4 streptomyce
13	42.5	46.2	1584	5 Q9NKB5	Q9NKB5 drosophila
14	42	45.7	338	5 Q9NNP0	Q9NNP0 leishmania
15	41.5	45.1	475	8 Q9THU1	Q9THU1 kalmia poli
16	41	44.6	98	2 Q9RXL0	Q9RXL0 deinococcus
17	41	44.6	246	10 Q81865	Q81865 arabidopsis
18	41	44.6	376	2 Q9X659	Q9X659 streptomyce
19	41	44.6	534	5 Q16636	Q16636 caenorhabdi

20	41	44.6	557	10 Q9W3C5	Q9W3C5 arabidopsis
21	41	44.6	587	12 Q68518	Q68518 hepatitis c
22	41	44.6	681	12 Q68519	Q68519 hepatitis c
23	41	44.6	1330	10 Q9L160	Q9L160 oryza sativ
24	41	44.6	7257	2 Q9K127	Q9K127 sorangium c
25	40.5	44.0	648	5 Q9V321	Q9V321 drosophila
26	40	43.5	57	12 Q9WGF1	Q9WGF1 tt virus. h
27	40	43.5	57	12 Q9QUF2	Q9QUF2 tt virus. h
28	40	43.5	71	13 Q9O659	Q9O659 gallus gall
29	40	43.5	94	10 Q9LWS4	Q9LWS4 oryza sativ
30	40	43.5	397	2 Q85746	Q85746 klebsiella
31	40	43.5	416	5 Q16868	Q16868 drosophila
32	40	43.5	419	9 Q38619	Q38619 xanthomonas
33	40	43.5	439	2 P74033	P74033 synechocyst
34	40	43.5	446	2 Q9RD00	Q9RD00 streptomyce
35	40	43.5	507	5 Q9W477	Q9W477 drosophila
36	40	43.5	659	5 Q94600	Q94600 leishmania
37	40	43.5	783	12 Q68948	Q68948 hepatitis c
38	40	43.5	832	5 Q16360	Q16360 caenorhabdi
39	40	43.5	2187	11 P70670	P70670 mus musculu
40	40	43.5	3010	12 Q68949	Q68949 hepatitis c
41	40	43.5	3010	12 Q93077	Q93077 hepatitis c
42	40	43.5	3012	12 Q9WIK7	Q9WIK7 hepatitis c
43	40	43.5	3014	12 Q86614	Q86614 hepatitis c
44	40	43.5	4124	5 Q97218	Q97218 leishmania
45	39.5	42.9	136	4 Q9UIG4	Q9UIG4 homo sapien

ALIGNMENTS

RESULT 1					
Q95407	095407	PRELIMINARY;	PRT;	300 AA.	
ID	095407;				
AC	095407;				
DT	01-MAY-1999 (TReMBLrel. 10, Created)				
DT	01-MAY-1999 (TReMBLrel. 10, Last sequence update)				
DT	01-OCT-2000 (TReMBLrel. 15, Last annotation update)				
DE	DECOY RECEPTOR 3 (M68) (M68C) (M68E).				
GN	DCR3 OR TR6.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=99087326; PubMed=9872321;				
RA	Pitti R.M., Marsters S.A., Lawrence D.A., Roy M., Kischkel F.C.,				
RA	Dowd P., Huang A., Donahue C.J., Sherwood S.W., Baldwin D.T.,				
RA	Godowski P.J., Wood W.I., Gurney A.L., Hillan K.J., Cohen R.L.,				
RA	Goddard A.D., Botstein D., Ashkenazi A.				
RT	"Genomic amplification of a decoy receptor for Fas ligand in lung and				
RT	colon cancer."				
RL	Nature 396:699-703(1998).				
[2]					
RN	SEQUENCE FROM N.A.				
RP	TISSUE=BL00D;				
RC	MEDLINE=99253915; PubMed=10318773;				
RA	Yu K.Y., Kwon B., Ni J., Zhai Y., Ebner R., Kwon B.S.;				
RT	"A newly identified member of tumor necrosis factor receptor				
RT	superfamily (TR6) suppresses LIGHT-mediated apoptosis."				
RL	J. Biol. Chem. 274:13733-13736(1999).				
[3]					
RN	SEQUENCE FROM N.A.				
RC	TISSUE=PANCREAS;				
RX	MEDLINE=20122600; PubMed=10655513;				
RA	Bai C., Connolly B., Metzker M.L., Hilliard C.A., Liu X., Sandig V.,				
RA	Soderman A., Galloway S.M., Liu O., Austin C.P., Caskey C.T.;				
RT	"Overexpression of M68/Dcr3 in human gastrointestinal tract tumors				
RT	independent of gene amplification and its location in a four-gene				
RT	cluster."				
RL	Proc. Natl. Acad. Sci. U.S.A. 97:1230-1235(2000).				
DR	EMBL; AF104419; AAD03056.1; -.				

```

DR EMBL; AF134240; AAD29688.1; -
DR EMBL; AF217796; AAF35244.1; -
DR EMBL; AF217793; AAF33685.1; -
DR EMBL; AF217794; AAF33686.1; -
DR HSSP; P25942; 1CDF.
DR INTERPRO; IPR000561; -
DR INTERPRO; IPR001368; -
DR PFAM; PF00020; TNFR_C6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -; 1.
KW Receptor.
SQ SEQUENCE 300 AA; 32679 MW; F90AEE33718449AF CRC64;

Query Match 100.0%; Score 92; DB 4; Length 300;
Best Local Similarity 100.0%; Pred. No. 5.2e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AETPTYPWRDAETGER 16
| | | | | | | | | | | | | | | |
DB 31 AETPTYPWRDAETGER 46

RESULT 2
O42801 PRELIMINARY; PRT; 537 AA.
AC 042801;
DT 01-JUN-1998 (TRENBLrel. 06, Created)
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
DE SUCROSE: SUCROSE 1-FRUCTOSYLTRANSFERASE PRECURSOR (EC 3.2.1.26).
OS Aspergillus foetidus.
OC Eukaryota; Fungi; Ascomycota; Eurotiales; Trichocomaceae;
OC anamorphic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=63131;
RN [1]
RP MEDLINE=98155160; PubMed=9495772;
RA Rehm J., Willmitzer L., Heyer A.G.;
RT Production of 1-ketose in transgenic yeast expressing a
fructosyltransferase from Aspergillus foetidus.;
RL J. Bacteriol. 180:1305-1310(1998).
CC -1- SIMILARITY: BELONGS TO FAMILY 32 OF GLYCOSYL HYDROLASES.
DR EMBL; AJ000493; CA04131.1; -
DR INTERPRO; IPR001362; -
DR PFAM; PF00251; Glyco_hydro_32; 2.
DR PROSITE; PS00609; GLYCOSYL_HYDROL_F32; 1.
KW Signal; Transferase; Hydrolase; Glycosidase.
FT SIGNAL 19 POTENTIAL.
FT CHAIN 20 537 SUCROSE: SUCROSE 1-FRUCTOSYLTRANSFERASE.
SQ SEQUENCE 537 AA; 59150 MW; 7483D0530EA35671 CRC64;

Query Match 53.3%; Score 49; DB 3; Length 537;
Best Local Similarity 72.7%; Pred. No. 7;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AETPTYPWRDA 11
| | | | | | | |
DB 339 ANIPTYPWRS 349

RESULT 3
P91905 PRELIMINARY; PRT; 5071 AA.
AC P91905;
DT 01-MAY-1997 (TRENBLrel. 03, Created)
DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE RYANODINE RECEPTOR.
GN RYR-1.

```

```

OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Sakube Y., Ando H., Kagawa H.;
RT "An abnormal ketamine response in mutants defective in the ryanodine
receptor gene, ryr-1(nuc-68) of Caenorhabditis elegans.";
RL J. Mol. Biol. 0:0-0(1997).
DR EMBL; D45899; BAA08309.1; -
DR INTERPRO; IPR000107; -
DR INTERPRO; IPR000561; -
DR INTERPRO; IPR000699; -
DR INTERPRO; IPR001215; -
DR INTERPRO; IPR001682; -
DR INTERPRO; IPR002048; -
DR INTERPRO; IPR002202; -
DR INTERPRO; IPR003032; -
DR PFAM; PF00036; efhand; 1.
DR PFAM; PF00622; SPRY; 3.
DR PFAM; PF01365; KYDR_IIPR; 1.
DR PFAM; PF02026; KYR; 4.
DR PRINTS; PR00795; RYANODINER;
DR PROSITE; PS00018; EF_HOAND; UNKNOWN_2.
DR PROSITE; PS00066; HMG_COA_REDUCTASE_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
SQ SEQUENCE 5071 AA; 576357 MW; 145EF32288613324 CRC64;

Query Match 49.5%; Score 45.5; DB 5; Length 5071;
Best Local Similarity 71.4%; Pred. No. 2.7e+02;
Matches 10; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 2 EPTYPWRDAETGE 15
| | | | | | | | | | | |
DB 2507 ENPTTPSKDAE-GE 2519

RESULT 4
O94279 PRELIMINARY; PRT; 5107 AA.
AC 094279;
DT 01-FEB-1997 (TRENBLrel. 02, Created)
DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE PARTIAL CDS.
DE UNC-68.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RC MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fultson L.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkneen R.,
RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans.";
RT Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;

```


PA (BIOJ) BIOGEN INC.
 XX
 PI Tschoep J;
 XX
 CC WPI; 1999-347693/29.
 DR N-PSDB; X76052.
 XX
 CC New tumour necrosis factor family receptor OPG-2
 PT
 XX
 CC Claim 1; Page 18; 22pp; English.
 PS
 CC The present sequence represents a mammalian tumour necrosis factor
 CC receptor, designated OPG-2. OPG-2, is a member of the tumour necrosis
 CC factor receptor family, and can be used: (i) to raise specific
 CC antibodies (Ab), (ii) to treat osteopenic disorders associated with
 CC excessive osteoclast activity, e.g. primary osteoporosis, Paget's
 CC disease, hyperglycaemia of malignancy, or osteolytic metastases; (iii)
 CC for affinity purification of cognate ligands, and (iv) to screen for
 CC ligands (antagonists or agonists). Ab, or other OPG-2 blocking agents
 CC such as soluble forms of the protein, are used to prevent, or reduce
 CC severity of, an immune response, and for treating cancer. They can also
 CC be used in diagnostic assays. The nucleic acid sequence encoding OPG-2
 CC can be used as a probe to isolate related sequences from other species.
 XX
 SQ Sequence 300 AA;
 Query Match 100.0%; Score 124; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 5.5e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ERCRYCNVLCGEREEERACH 21
 Db |||||
 86 ercrynvlcgerereearach 106
 RESULT 12
 Y06817
 ID Y06817 standard; Protein; 300 AA.
 AC Y06817;
 DT 24-JUN-1999 (first entry)
 DE Human Dcr3 polypeptide.
 XX
 KW Dcr3 polypeptide; tumour necrosis factor receptor; TNFR; Fas ligand;
 KW apoptosis; T cell mediated immune response; allergy; asthma; cancer;
 KW rheumatoid arthritis; Crohn's disease; guest vs. host disease; human;
 KW gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN W09914330-A1.
 XX
 PD 25-MAR-1999.
 XX
 PF 18-SEP-1998; 98WO-US19661.
 XX
 PR 30-JUL-1998; 98US-0094640.
 PR 18-SEP-1997; 97US-0059288.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Ashkenazi AJ, Botstein D, Dodge KH, Goddard A, Gurney AL;
 PI Kim KJ, Lawrence DA, Pitti R, Roy MA, Tumas DB;
 PI Wood WI;
 XX
 DR WPI; 1999-244032/20.
 DR N-PSDB; X32744.
 XX
 PT Dcr3 polypeptide related to tumor necrosis factor receptor
 XX

PS Claim 5; Fig 1; 88pp; English.
 XX
 CC This represents a human Dcr3 polypeptide, a homologue of tumour necrosis
 CC factor receptor (TNFR) polypeptide. Host cells containing a vector
 CC comprising the Dcr3 nucleic acid can be used for the recombinant
 CC expression of the protein. Dcr3 binds to Fas ligand, so it (or its
 CC chimeras) are useful for modulating apoptosis in mammalian cells, also
 CC other Fas-ligand induced activities, particularly to inhibit T cell
 CC mediated immune responses, e.g. in treatment of allergy, asthma,
 CC rheumatoid arthritis, Crohn's disease, guest vs. host disease etc. Dcr3
 CC may also be used to identify specific binding proteins, potential
 CC inhibitors. Antibodies against Dcr3 are used to treat cancer,
 CC specifically of the lung and colon, also in diagnosis and for affinity
 CC purification of the protein. Detecting mutations in the gene for Dcr3 is
 CC also used to diagnose cancer, or predisposition to it. Dcr3 nucleic acid
 CC is useful as hybridization probe to detect genomic or related sequences;
 CC for chromosome and gene mapping; as source of antisense sequences;
 CC expression of recombinant Dcr3 and to generate transgenic animals (for
 CC development and screening of therapeutic agents), also for in vivo or
 CC ex vivo gene therapy.
 XX
 SQ Sequence 300 AA;
 Query Match 100.0%; Score 124; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 5.5e-09;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ERCRYCNVLCGEREEERACH 21
 Db |||||
 86 ercrynvlcgerereearach 106
 RESULT 13
 W97749
 ID W97749 standard; Protein; 300 AA.
 XX
 AC W97749;
 XX
 DT 21-MAY-1999 (first entry)
 DE Human tumour necrosis factor receptor ZTNFR-5.
 XX
 KW ZTNFR-5; tumour necrosis factor receptor; TNFR; human;
 KW cell maturation; bone cell regulation.
 XX
 OS Homo sapiens.
 XX
 PN Key Location/Qualifiers
 FH Peptide 1..23
 FT /note= "signal peptide"
 FT Protein 24..300
 FT /note= "mature protein"
 FT Domain 24..194
 FT /note= "extracellular domain"
 FT Region 49..71
 FT /note= "cysteine-rich pseudo-repeat 1"
 FT Region 72..113
 FT /note= "cysteine-rich pseudo-repeat 1"
 FT Region 114..151
 FT /note= "cysteine-rich pseudo-repeat 1"
 FT Region 152..194
 FT /note= "cysteine-rich pseudo-repeat 1"
 XX
 PN W09904001-A1.
 XX
 PD 28-JAN-1999.
 XX
 PF 21-JUL-1998; 98WO-US15072.
 XX
 PR 21-JUL-1997; 97US-0053203.
 XX
 PA (ZYMO) ZYMOGENETICS INC.

XX Farrah TM;
PI
XX
DR WPI; 1999-132245/11.
DR N-PSDB; X07226.
XX
XX Novel tumour necrosis factor receptor zTNFR5 - useful for
PT regulating maturation of TNF-ligand bearing cells
XX
XX
PS Claim 1; Page 84-85; 109pp; English.
XX
XX This polypeptide comprises a new, secreted tumour necrosis factor
CC receptor (see W97749), designated zTNFR-5. Novel zTNFR-5 encoding
CC polynucleotides and polypeptides were initially identified by
CC querying an expressed sequence tag (EST) database for sequences
CC homologous to conserved motifs within the TNF receptor family.
CC Based on this search, a contig of 16 ESTs (see X07226) was
CC constructed. zTNFR-5 polypeptides comprise 4 cysteine-rich repeats
CC (see also W97750-55) that are homologous to other TNF receptors, in
CC particular the soluble, secreted TNF receptor osteoprotegerin.
CC zTNFR-5 polypeptide can be prepared by recombinant methods. The
CC polypeptide, especially the extracellular domain, can be used to
CC generate a soluble variant of zTNFR-5. The polypeptides and
CC nucleic acids can be used to screen for ligands, agonists and
CC antagonists of zTNFR-5. The polypeptides can be used in bone cell
CC regulation and to regulate the maturation of TNF ligand-bearing
CC cells such as T- or B-cells, lymphocytes, peripheral blood
CC mononuclear cells, polymorphonuclear leukocytes, fibroblasts or
CC haematopoietic cells.
XX
SQ Sequence 300 AA;

Query Match 100.0%; Score 124; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 5.5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERCRYCNVLCGEREEARACH 21
Db 86 ercrycnvlgereearach 106
|||||

RESULT 15
Y42183 ID Y42183 standard; Protein; 302 AA.
XX
AC Y42183;
XX
DT 17-DEC-1999 (first entry)
XX
DE Human FLINT #2 protein sequence.
XX
KW Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
KW reperfusion-associated injury; aplastic anaemia; differentiation;
KW growth; myelodysplastic syndrome; pancytopenic condition;
KW myocardial ischaemia.
XX
OS Homo sapiens.
XX
PN W09950413-A2.
XX
PD 07-OCT-1999.
XX
PF 30-MAR-1999; 99WO-US06797;
XX
PR 30-MAR-1998; 98US-0079856.
PR 20-MAY-1998; 98US-0086074.
PR 09-SEP-1998; 98US-0099643.
PR 17-DEC-1998; 98US-0112577.
PR 18-DEC-1998; 98US-0112703.
PR 18-DEC-1998; 98US-0112933.
PR 22-DEC-1998; 98US-0113407.
XX
PA (ELIL) LILLY & CO ELI.
XX
PI Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
PI Song HY, Wang J, Wu X, Zuckerman SH;
XX
XX WPI; 1999-591319/50.
DR N-PSDB; Z25376.
XX
XX Use of mature FLINT for treating acute liver failure, inflammation,
PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic

XX
PI
XX
DR WPI; 1999-132245/11.
DR N-PSDB; X07226.
XX
XX Novel tumour necrosis factor receptor zTNFR5 - useful for
PT regulating maturation of TNF-ligand bearing cells
XX
XX
PS Claim 1; Page 84-85; 109pp; English.
XX
XX This polypeptide comprises a new, secreted tumour necrosis factor
CC receptor (see W97749), designated zTNFR-5. Novel zTNFR-5 encoding
CC polynucleotides and polypeptides were initially identified by
CC querying an expressed sequence tag (EST) database for sequences
CC homologous to conserved motifs within the TNF receptor family.
CC Based on this search, a contig of 16 ESTs (see X07226) was
CC constructed. zTNFR-5 polypeptides comprise 4 cysteine-rich repeats
CC (see also W97750-55) that are homologous to other TNF receptors, in
CC particular the soluble, secreted TNF receptor osteoprotegerin.
CC zTNFR-5 polypeptide can be prepared by recombinant methods. The
CC polypeptide, especially the extracellular domain, can be used to
CC generate a soluble variant of zTNFR-5. The polypeptides and
CC nucleic acids can be used to screen for ligands, agonists and
CC antagonists of zTNFR-5. The polypeptides can be used in bone cell
CC regulation and to regulate the maturation of TNF ligand-bearing
CC cells such as T- or B-cells, lymphocytes, peripheral blood
CC mononuclear cells, polymorphonuclear leukocytes, fibroblasts or
CC haematopoietic cells.
XX
SQ Sequence 300 AA;

Query Match 100.0%; Score 124; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 5.5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERCRYCNVLCGEREEARACH 21
Db 86 ercrycnvlgereearach 106
|||||

RESULT 14
W95082 ID W95082 standard; Protein; 300 AA.
XX
AC W95082;
XX
DT 20-MAY-1999 (first entry)
XX
DE Orphan receptor (HUMAN NTR-1) polypeptide.
XX
KW HUMAN NTR-1; orphan receptor; osteoprotegerin; OPG; TNFR; human;
KW tumour necrosis factor receptor; muscle disorder; bone mass; screening;
KW muscle metabolism; binding agent; cognate ligand.
XX
OS Homo sapiens.
XX
PN W09907738-A2.
XX
PD 18-FEB-1999.
XX
PF 04-AUG-1998; 98WO-US16202.
XX
PR 06-AUG-1997; 97US-0054869.
XX
PA (PROC) PROCTER & GAMBLE CO.
XX (REGE-) REGENERON PHARM INC.
XX
PI Maslakowski PJ, Morris J, Valenzuela DM;
XX
XX WPI; 1999-167365/14.
DR N-PSDB; X22300.

PT and proinflammatory activity
XX
PS Example 2; Fig 2; 99pp; English.
XX
CC The present invention describes therapeutic applications of mature FLINT
CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
CC (mFLINT), which is a member of the tumour necrosis factor receptor
CC superfamily, is used for treating acute liver failure, inflammation of
CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
CC injury or disorder such as hypercoagulation (including use with
CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
CC or disorder, Type I diabetes, cancer, cell damage or damage to an
CC innocent bystander tissue that is induced by a chemotherapeutic agent or
CC therapeutic irradiation, treating haematopoietic progenitor cells that
CC have been exposed to therapeutic radiation or chemotherapy, aplastic
CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
CC also used for promoting the growth or differentiation of a haematopoietic
CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
CC resulting from abnormal myocardial ischaemia. The present sequence
XX represents human FLINT.
SQ Sequence 302 AA;

Query Match 100.0%; Score 124; DB 20; Length 302;
Best Local Similarity 100.0%; Pred. No. 5.5e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERCRYCNVLCGEREEARACH 21
Db ||||||||||||||||
86 ercrycnvlgereeeatrch 106

Search completed: January 30, 2001, 16:45:30
Job time: 590 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:50:34 ; Search time 149.64 Seconds
(without alignments)
9.529 Million cell updates/sec

Title: US-09-518-931-4_COPY_86_106

Perfect score: 124

Sequence: 1 ECRYCNVLCGEREEARACH 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR66: *
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	54	43.5	477	2 S23257	alpha-amylase (EC
2	52.5	42.3	318	2 JC4963	metalloproteinase
3	51	41.1	319	2 A53502	follicistatin - Atri
4	51	41.1	469	2 G69267	conserved hypothet
5	50.5	40.7	2529	2 A56923	transcription fact
6	50	40.3	182	2 T46387	hypothetical prote
7	48	38.7	62	2 S62863	toxin IV-5 - Tityu
8	48	38.7	272	2 T20991	hypothetical prote
9	48	38.7	657	2 B82176	helicase-related p
10	47	37.9	114	2 E64467	hypothetical prote
11	47	37.9	344	2 A27701	follicistatin precu
12	47	37.9	377	2 T43739	heat shock protein
13	47	37.9	445	2 T33986	hypothetical prote
14	46.5	37.5	313	2 T26308	hypothetical prote
15	46.5	37.5	606	2 A54665	netrin-1 precursor
16	46	37.1	63	2 A23727	neurotoxin V - bar
17	46	37.1	336	2 A72322	hypothetical prote
18	46	37.1	337	2 T47079	follicistatin - shee
19	46	37.1	344	2 T45894	follicistatin - bovi
20	46	37.1	425	2 JC5909	A333 protein - fru
21	46	37.1	488	2 T32754	hypothetical prote
22	45	36.3	343	2 S55369	follicistatin - chic
23	45	36.3	348	2 T28623	hypothetical prote
24	45	36.3	349	2 D72175	G2R protein - vari
25	45	36.3	349	2 D36858	gene G4R protein -
26	45	36.3	384	2 E70036	capsular polysacch
27	45	36.3	418	2 T41087	hypothetical prote
28	44.5	35.9	539	2 T49065	lymphoid-restrict
29	44.5	35.9	608	2 JQ1462	phosphoenolpyruvat

30	44	35.5	379	2 T04030	hypothetical prote
31	44	35.5	421	2 T20315	hypothetical prote
32	44	35.5	484	2 S44739	C02C2.1 protein -
33	44	35.5	573	2 T48584	auxin reponsive-li
34	44	35.5	576	2 T48585	auxin reponsive-li
35	44	35.5	3020	2 A43932	mucin 2 precursor,
36	43.5	35.1	170	2 A83135	4-hydroxyphenylace
37	43.5	35.1	1548	2 T25808	hypothetical prote
38	43	34.7	66	1 NTSR9E	neurotoxin M9 - le
39	43	34.7	98	1 A27510	ferredoxin 2[4Fe-4
40	43	34.7	331	2 T01838	hypothetical prote
41	43	34.7	348	1 JQ0701	methane monooxygen
42	43	34.7	379	2 E59332	heterodisulfide re
43	43	34.7	432	2 T35527	hypothetical prote
44	43	34.7	480	2 B64308	hypothetical prote
45	43	34.7	575	2 T48579	auxin-responsive-1

ALIGNMENTS

RESULT 1

S23257
alpha-amylase (EC 3.2.1.1) precursor - Alteromonas haloplanktis
C:Species: Alteromonas haloplanktis
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999
C:Accession: S23257
R:Feller, G.; Lonhienne, T.; Deroanne, C.; Libioulle, C.; van Beeumen, J.; Gerday, C.
J. Biol. Chem. 267, 5217-5221, 1992
A:Title: Purification, characterization, and nucleotide sequence of the thermolabile
A:Reference number: S23257; MUID:92184768
A:Accession: S23257
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-477 <FEL>
A:Cross-references: EMBL:X58627; NID:g38819; PID:g38820
C:Function:
A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds
A:Pathway: glycogen/starch degradation
C:Superfamily: mammalian alpha-amylase; alpha-amylase core homology
C:Keywords: glycosidase; hydrolase; polysaccharide degradation
F:166-291/Domain: alpha-amylase core homology <AMY>

Query Match 43.5%; Score 54; DB 2; Length 477;
Best Local Similarity 52.9%; Pred. No. 3.9;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 4 RVCNVLCGEREEARAC 20

Db 424 QYCNVLKGLSADAKSC 440

RESULT 2

JC4963
metalloproteinase 1 (EC 3.4.24.-) - human
N:Alternate names: PRSM1 metalloproteinase
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1996 #sequence_revision 31-Dec-1996 #text_change 07-Dec-1999
C:Accession: JC4963; G02657
R:Scott, I.C.; Halliia, R.; Jenkins, J.M.; Mehan, S.; Apostolou, S.; Winqvist, R.; Cal
Gene 174, 135-143, 1996
A:Title: Molecular cloning, expression and chromosomal localization of a human gene e
A:Reference number: JC4963; MUID:97017139
A:Accession: JC4963
A:Molecule type: mRNA
A:Residues: 1-318 <SCO>
A:Cross-references: EMBL:U58048; NID:g1354930; PIDN:AAC50775.1; PID:g1354931
C:Experimental source: placental cell
C:Genetics:
A:Gene: prsm1
A:Map position: 16q24.3
C:Keywords: hydrolase; metalloproteinase

conserved hypothetical protein AF0143 - *Archaeoglobus fulgidus*
C:Species: *Archaeoglobus fulgidus*
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: G69267
R:Klenk, H. P.; Clayton, R. A.; Tomb, J. F.; White, O.; Nelson, K. E.; Ketchum, K. A.; Dodson
.: Fleischmann, R. D.; Quackenbush, J.; Lee, N. H.; Sutton, G. G.; Gill, S.; Kirkness, E. F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J. D.; Weidman, J. F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M. D.; Spriggs, T.; Artiaich, P.; Kaine, B. P.; Sykes, S.
Smith, H. O.; Woese, C. R.; Venter, J. C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A:Reference number: A69250; MUID:98049343
A:Accession: G69267

QY 3 CRYCNVLCGEREEARAC 20
| : | | : | | | : |
Db 165 CKRCKVMLGETVSSARSC 182

```
RESULT 7
S62863
toxin IV-5 - Tityus bahiensis
C:Species: Tityus bahiensis
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 07-May-1999
C:Accession: S62863
R:Beccerilli, B.; Corona, M.; Coronas, F.I.V.; Zamudio, F.; Calderon-Aranda, E.S.; Fletcher
Biochem. J. 313, 753-760, 1996
A:Title: Toxic peptides and genes encoding toxin gamma of the Brazilian scorpions Tityus
A:Reference number: S62861; MUID:96190713
A:Accession: S62863
A:Molecule type: protein
A:Residues: 1-62 <BEC>
A:Superfamily: scorpion neurotoxin
C:Keywords: neurotoxin; venom
F:12-62,16-38,24-45,28-47/Disulfide bonds: #status predicted

Query Match 38.7%; Score 48; DB 2; Length 62;
Best Local Similarity 35.3%; Pred. No. 5.2;
Matches 6; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Qy 5 YCNVLCGEREEERACH 21
Db 23 YCDKLCGDKKADSGYCY 39

RESULT 8
T20991
hypothetical protein F15G9.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T20991
R:Stuston, J.
submitted to the EMBL Data Library, December 1994
A:Reference number: Z19355
A:Accession: T20991
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-272 <WIL>
A:Cross-references: EMBL:Z47068; PIDN:CAA87334.2; GSPDB:GN00028; CESP:F15G9.5
A:Experimental source: clone F15G9
C:Genetics:
A:Gene: CESP:F15G9.5
A:Map position: X
A:Introns: 55/2; 99/2; 136/1; 156/2; 206/3

Query Match 38.7%; Score 48; DB 2; Length 272;
Best Local Similarity 38.1%; Pred. No. 17;
Matches 8; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

Qy 1 ERRCRYNVLCGEREEERACH 21
Db 73 ERDNYCYLCCGSEHQCAAH 93

RESULT 9
B82176
helicase-related protein VC1636 [imported] - Vibrio cholerae (group O1 strain N16961)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Sep-2000
C:Accession: B82176
R:Heidelberger, J.F.; Eelsen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B.
I., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A62035; MUID:20406833
A:Accession: B82176
A>Status: preliminary
A:Molecule type: DNA

A:Residues: 1-657 <HEI>
A:Cross-references: GB:AE004241; GB:AE003852; NID:g9656142; PIDN:AAF94787.1; GSPDB:GN
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1636
A:Map position: 1

Query Match 38.7%; Score 48; DB 2; Length 657;
Best Local Similarity 40.0%; Pred. No. 33;
Matches 8; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 2 RCRYCNVLCGEREEERACH 21
Db 493 RAKYCNECGADNDIAARICH 512

RESULT 10
E64467
hypothetical protein MJ1342 - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C:Accession: E64467
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannasc
A:Reference number: A64300; MUID:96337999
A:Accession: E64467
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-114 <BUL>
A:Cross-references: GB:U67574; GB:L77117; NID:g1591978; PIDN:AAB99359.1; PID:g1591984
C:Genetics:
A:Map position: FOR1289459-1289803
C:Superfamily: Methanococcus jannaschii hypothetical protein MJ1342

Query Match 37.9%; Score 47; DB 2; Length 114;
Best Local Similarity 40.0%; Pred. No. 11;
Matches 8; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 1 ERRCRYNVLCGEREEERAC 20
Db 54 DKCRYCPKOIGAEIPGSAC 73

RESULT 11
A27701
follistatin precursor - pig
N:Contains: follistatin short form precursor
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A27701; A40064; B40064; A39969; A47139
R:Shimasaki, S.; Koga, M.; Esch, F.; Mercado, M.; Cooksey, K.; Koba, A.; Ling, N.
Biochem. Biophys. Res. Commun. 152, 717-723, 1988
A:Title: Porcine follistatin gene structure supports two forms of mature follistatin
A:Reference number: A27701; MUID:88209050
A:Accession: A27701
A:Molecule type: DNA
A:Residues: 1-344 <SHI>
A:Cross-references: GB:M19529; NID:g164458; PIDN:AAA31036.1; PID:g164459
R:Esch, F.S.; Shimasaki, S.; Mercado, M.; Cooksey, K.; Ling, N.; Ying, S.; Ueno, N.;
Mol. Endocrinol. 1, 849-855, 1987
A:Title: Structural characterization of follistatin: a novel follicle-stimulating hor
A:Reference number: A40064; MUID:91042571
A:Accession: A40064
A:Molecule type: mRNA
A:Residues: 1-344 <ESC1>
A:Cross-references: GB:M36512; GB:M36513; NID:g164461
A:Accession: B40064
```

```

RESULT 13
T33986
hypothetical protein T12B3.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T33986
R:Nelson, J.; Wohlmann, P.; Antoniou, B.; Fulton, B.
submitted to the EMBL Data Library, February 1999
A:Description: The sequence of C. elegans cosmid T12B3.
A:Reference number: Z21450
A:Accession: T33986
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-446 <NLE>
A:Cross-references: EMBL:AF125960; PTDN:AAD14737.1; GSPDB:GN00022; CESP:T12B3.1
A:Experimental source: strain Bristol N2; clone T12B3
C:Genetics:
A:Gene: CESP:T12B3.1
A:Map position: 4
A:Introns: 21/1; 100/2; 139/1; 228/3; 271/3; 351/3; 391/3

Query Match 37.9%; Score 47; DB 2; Length 446;
Best Local Similarity 50.0%; Pred. No. 33;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

.Qy 2 RCRCNVLCGERESEA 17
|||||:|:|:|:|:|
Db 50 RCRCIVVVGQTEIQA 65

RESULT 14
T26308
hypothetical protein W09D6.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T26308
R:Gardner, A.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z20195
A:Accession: T26308
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-313 <WII>
A:Cross-references: EMBL:Z82078; PTDN:CAB04942.1; GSPDB:GN00021; CESP:W09D6.2
A:Experimental source: clone W09D6
C:Genetics:
A:Gene: CESP:W09D6.2
A:Map position: 3
A:Introns: 61/3; 91/1; 117/2; 162/2; 202/1; 306/2

Query Match 37.5%; Score 46.5; DB 2; Length 313;
Best Local Similarity 38.7%; Pred. No. 30;
Matches 12; Conservative 1; Mismatches 5; Indels 13; Gaps 2

Qy 3 CR-----YCNVLCGE-----REEEARAC 20
||| |||||:|
Db 274 CRNSELDCNVSCAEPDIADYIYRELPESC 304

RESULT 15
A54665
netrin-1 precursor - chicken
C:Species: Gallus gallus (chicken)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 21-Jul-2000
C:Accession: A54665
R:Serafini, T.; Kennedy, T.E.; Galko, M.J.; Mirzayan, C.; Jessell, T.M.; Tessier-Lavigne, M.
Cell 78, 409-424, 1994
A:Title: The netrins define a family of axon outgrowth-promoting proteins homologous to Drosophila netrin
A:Reference number: A54665; MUID:94340732
A:Accession: A54665

```


A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-606 <SER>
A:Cross-references: GR:L34549: NID:g529418: PIDN:AAA60369.1; PID:g529419
C:Superfamily: laminin-type EGF-like homology
F:287-340/Domain: laminin-type EGF-like homology <LEG>

Query Match 37.5%; Score 46.5; DB 2; Length 606;
Best Local Similarity 50.0%; Pred. No. 50;
Matches 9; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

OY 4 RYCNVLCGEREEARACH 21
||| | : || : ||
Db 78 RYC-VVTEKGEQVRSCH 94

Search completed: January 30, 2001, 16:50:36
Job time: 726 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 17:06:09 ; Search time 80.04 seconds
(without alignments)
8.473 Million cell updates/sec

Title: US-09-518-931-4_COPY_86_106
Perfect score: 124
Sequence: 1 ERCRYCNVLCGEREEARACH 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	54	43.5	669	1 AMY_ALTHA	P29957 alteromonas
2	48	38.7	62	1 SCX1_TITBA	P56608 titysus bahi
3	47	37.9	344	1 FSA_PIG	P10669 sus scrofa
4	46.5	37.5	606	1 NET1_CHICK	Q09022 gallus gall
5	46	37.1	63	1 SCX5_GENSC	P46066 centruride
6	46	37.1	135	1 LSG7_RAT	P97590 rattus norv
7	46	37.1	337	1 FSA_SHEEP	P31514 ovis aries
8	46	37.1	344	1 FSA_BOVIN	P50291 bos taurus
9	45	36.3	349	1 VC22_VARV	P34015 variola vir
10	44.5	35.9	608	1 PPCK_NEOFR	P22130 neocallimas
11	44	35.5	344	1 FSA_HORSE	O62650 equus cabal
12	44	35.5	484	1 YKH1_CABEL	P34269 caenorhabdi
13	44	35.5	5179	1 MUC2_HUMAN	Q02817 homo sapien
14	43	34.7	66	1 SCX9_BUTEU	P09981 buthus eupe
15	43	34.7	98	1 FIXX_RHILT	P08710 rhizobium l
16	43	34.7	348	1 MEMC_METCA	P22868 methylococc
17	43	34.7	480	1 Y066_METUA	Q60377 methanococc
18	42.5	34.3	869	1 CLC6_HUMAN	P51797 homo sapien
19	42.5	34.3	870	1 CLC6_MOUSE	O35454 mus musculu
20	42	33.9	197	1 YRPA_SYNPZ	Q02426 synechococc
21	42	33.9	322	1 LIPA_RHIEI	O05941 rhizobium e
22	42	33.9	339	1 GRPH_HUMAN	Q13304 homo sapien
23	42	33.9	471	1 TNRL_BOVIN	O19131 bos taurus
24	42	33.9	621	1 FOR_THELI	Q56303 thermococcu
25	42	33.9	661	1 BAIH_EUBSP	P32370 eubacterium
26	42	33.9	1477	1 HPR7_HYDAT	Q25197 hydra atten
27	41.5	33.5	65	1 SCX3_MESMA	P15227 mesobuthus
28	41.5	33.5	84	1 SCXA_MESMA	O61705 mesobuthus
29	41.5	33.5	220	1 ANTA_HYDMA	P38977 hydra magni
30	41.5	33.5	703	1 EGFR_CHICK	P13387 gallus gall
31	41.5	33.5	3866	1 HRX_MOUSE	P55200 mus musculu
32	41.5	33.5	3969	1 HRX_HUMAN	Q03164 homo sapien
33	41	33.1	157	1 YSAA_ECOLI	P56256 escherichia

34	41	33.1	190	1 Y416_METJA	Q57859 methanococc
35	41	33.1	338	1 FSA_HUMAN	P19883 homo sapien
36	41	33.1	344	1 FSA_MOUSE	P47931 mus musculu
37	41	33.1	344	1 FSA_RAT	P21674 rattus norv
38	41	33.1	472	1 USP_CHOFU	O76202 chorisoneu
39	41	33.1	497	1 LMP2_EBV	P13285 epstein-bar
40	41	33.1	704	1 KPC1_CABEL	P34722 caenorhabdi
41	41	33.1	727	1 KPC1_HUMAN	P49711 homo sapien
42	41	33.1	728	1 CTCF_CHICK	Q08705 gallus gall
43	41	33.1	755	1 NAPA_DESDE	P81186 desulfovibr
44	41	33.1	1400	1 RON_HUMAN	Q04912 homo sapien
45	41	33.1	1437	1 VGLM_BUNGE	P12430 bunyavirus

ALIGNMENTS

RESULT 1
ID AMY_ALTHA STANDARD; PRT; 669 AA.
AC P29957;
DT 01-APR-1993 (Rel. 25, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ALPHA-AMYLASE PRECURSOR (EC 3.2.1.1) (1,4-ALPHA-D-GLUCAN
DE GLUCANOHYDROLASE).
GN AMY.
OS Alteromonas haloplanktis.
OC Bacteria; Proteobacteria; gamma subdivision; Alteromonadaceae;
OC Pseudoalteromonas.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 25-64 AND 471-477.
RC STRAIN=A23;
RX MEDLINE=92184768; PubMed=1544904;
RA Feller G., Lonhienne T., Deroanne C., Libioule C., van Beeumen J.,
RA Gerday C.;
RT "Purification, characterization, and nucleotide sequence of the
RT thermolabile alpha-amylase from the antarctic psychrotroph
RT Alteromonas haloplanktis A23".
RL J. Biol. Chem. 267:5217-5221(1992).
RN [2]
RP REVISIONS.
RX STRAIN=A23;
RX MEDLINE=98241570; PubMed=9575155;
RA Feller G., D'Amico S., Benotmane A.M., Joly F., van Beeumen J.,
RA Gerday C.;
RT "Characterization of the C-terminal propeptide involved in
RT bacterial wall spanning of alpha-amylase from the psychrophile
RT Alteromonas haloplanktis".
RL J. Biol. Chem. 273:12109-12115(1998).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 25-477.
RC STRAIN=A23;
RX MEDLINE=98200460; PubMed=9541387;
RA Aghajari N., Feller G., Gerday C., Haser R.;
RT "Crystal structures of the psychrophilic alpha-amylase from
RT Alteromonas haloplanktis in its native form and complexed with an
RT inhibitor".
RL Protein Sci. 7:564-572(1998).
CC -|- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC
CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.
CC -|- COPACITOR: ALPHA-AMYLASE BINDS THREE CALCIUM IONS PER MOLECULE
CC REQUIRED FOR ITS ACTIVITY.
CC -|- MISCELLANEOUS: THIS ENZYME IS THERMOLABILE.
CC -|- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

NET1_CHICK
ID NET1_CHICK STANDARD; PRT; 606 AA.
AC Q90922;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE NETRIN-1 PRECURSOR.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
CC [1]
RN SEQUENCE FROM N.A.
RP STRAIN=WHITE LEGHORN; TISSUE=EMBRYONIC BRAIN;
RX MEDLINE=94340732; PubMed=6062384;
RA Serafini T., Kennedy T.E., Galko M.J., Mirzayan C., Jessell T.M.,
RT Tessier-Lavigne M.;
RT "The netrins define a family of axon outgrowth-promoting proteins
homologous to C. elegans UNC-6";
RL Cell 78:409-424(1994).
CC -!- FUNCTION: NETRINS CONTROL GUIDANCE OF CNS COMMISSURAL AXONS AND
PERIPHERAL MOTOR AXONS. PROMOTES NEURITE OUTGROWTH FROM
COMMISSURAL AXONS BUT ACTS AS A CHEMOREPULSANT FOR TROCHLEAR MOTOR
AXONS. THESE EFFECTS ARE MEDIATED BY DISTINCT RECEPTORS.
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -!- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC -!- SIMILARITY: CONTAINS 3 LAMININ EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 C345C DOMAIN.
CC -!- SIMILARITY: TO CAENORHABDITIS ELEGANS UNC-6 PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L34549; AAA60369.1; -;
DR HSPSP; P02468; IKLO.
DR INTERPRO: IPR000561; -;
DR INTERPRO: IPR001134; -;
DR INTERPRO: IPR001886; -;
DR INTERPRO: IPR002049; -;
DR PFAM: PF01759; NTR; 1.
DR PFAM: PF00053; laminin_EGF; 3.
DR PFAM: PF00055; laminin_Nterm; 1.
DR PROSITE: PS00022; EGF_1; 2.
DR PROSITE: PS01248; LAMININ_TYPE_EGF; 3.
DR Glycoprotein; Extracellular matrix; Signal; Laminin EGF-like domain;
KW Repeat.
FT SIGNAL. 1 25 POTENTIAL.
FT CHAIN 26 606 NETRIN-1.
FT DOMAIN 26 286 LAMININ N-TERMINAL (DOMAIN VI).
FT DOMAIN 287 455 3 X LAMININ EGF-LIKE REPEATS (LAMININ
DOMAIN V-LIKE).
FT DOMAIN 287 342 LAMININ EGF-LIKE 1.
FT DOMAIN 343 405 LAMININ EGF-LIKE 2.
FT DOMAIN 406 455 LAMININ EGF-LIKE 3.
FT DOMAIN 456 606 C345C (DOMAIN C).
FT SITE 532 534 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 121 154 BY SIMILARITY.
FT DISULFID 287 296 BY SIMILARITY.
FT DISULFID 289 306 BY SIMILARITY.
FT DISULFID 308 317 BY SIMILARITY.
FT DISULFID 320 340 BY SIMILARITY.
FT DISULFID 343 352 BY SIMILARITY.
FT DISULFID 345 370 BY SIMILARITY.
FT DISULFID 373 382 BY SIMILARITY.
FT DISULFID 385 403 BY SIMILARITY.
FT DISULFID 406 418 BY SIMILARITY.
FT DISULFID 408 425 BY SIMILARITY.
FT DISULFID 427 436 BY SIMILARITY.

FT DISULFID 439 453 BY SIMILARITY.
FT DISULFID 478 546 BY SIMILARITY.
FT DISULFID 493 603 BY SIMILARITY.
FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 118 118 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 419 419 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 606 AA; 68126 MW; 9BF0E3B07A71AE1 CRC64;

Query Match 37.5%; Score 46.5; DB 1; Length 606;
Best Local Similarity 50.0%; Pred. No. 15;
Matches 9; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 4 RYCNVLCGEREEARACH 21
DB 78 RYC-VVTEKGEEVRSCH 94
||| | : : ||: ||: ||

RESULT 5
SCX5_CENS
ID SCX5_CENS STANDARD; PRT; 63 AA.
AC P46066;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE NEUROTOXIN V (CSE-V).
OS Centruroides sculpturatus (Bark scorpion).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC Buthoidea; Buthidae; Centruroides.
RN [1]
RP STRUCTURE BY NMR.
RC TISSUE=VENOM;
RX MEDLINE=95257401; PubMed=7739052;
RA Jablonsky M.J., Watt D.D., Krishna N.R.;
RT "Solution structure of an Old World-like neurotoxin from the venom of
the New World scorpion Centruroides sculpturatus Ewing.";
RL J. Mol. Biol. 248:449-458(1995).
CC -!- FUNCTION: BINDS TO SODIUM CHANNELS AND INHIBIT THE INACTIVATION OF
THE ACTIVATED CHANNELS, THEREBY BLOCKING NEURONAL TRANSMISSION.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC ALPHA-TOXIN SUBFAMILY.
DR PDB: 1NRA; 27-FEB-95.
DR PDB: 1NRB; 27-FEB-95.
DR INTERPRO: IPR002061; -;
DR PFAM: PF00537; toxin_3; 1.
DR PRINTS: PR00285; SCORPNTOXIN.
DR Neurotoxin; Sodium channel inhibitor; 3D-structure.
FT DISULFID 12 60
FT DISULFID 16 36
FT DISULFID 22 43
FT DISULFID 26 45
SQ SEQUENCE 63 AA; 7076 MW; 8BFEB36915AD8467 CRC64;

Query Match 37.1%; Score 46; DB 1; Length 63;
Best Local Similarity 47.1%; Pred. No. 2.3;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 5 YCNVLCGEREEARACH 21
DB 21 YCNCLCLERKADKGYCY 37
||| | : : ||: ||: ||

RESULT 6
LEG7_RAT
ID LEG7_RAT STANDARD; PRT; 135 AA.
AC P97590; O54958;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GALECTIN-7.


```
FT SIGNAL 1 29 POTENTIAL.
FT CHAIN 30 344 FOLLISTATIN.
FT CARBOHYD 124 124 N-LINKED (GLCNAC: . . .) (POTENTIAL).
FT CARBOHYD 288 288 N-LINKED (GLCNAC: . . .) (POTENTIAL).
SQ SEQUENCE 344 AA; 37959 MW; 018FEE8A84F97710 CRC64;

Query Match 37.1%; Score 46; DB 1; Length 344;
Best Local Similarity 42.1%; Pred. No. 11;
Matches 8; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

Qy 2 RCRYCNVLCGEREEARAC 20
   ||| | | | | |
Db 266 RCLSGELCPESKSEPC 284

RESULT 9
VC22_VARV STANDARD; PRT; 349 AA.
AC P34015;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE PROTEIN C22/B28 HOMOLOG.
GN G4R.
OS Variola virus.
OS Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
CC [1]
CC SEQUENCE FROM N.A.
RC STRAIN=INDIA-1967 / ISOLATE IND3;
RX MEDLINE=93202281; PubMed=8384129;
RA Shchelkunov S.N., Blinov V.M., Sandakhchiev L.S.;
RT "Genes of variola and vaccinia viruses necessary to overcome the host
   protective mechanisms.";
RL FEBS Lett. 319:80-83(1993).
CC -1- SIMILARITY: CONTAINS 2 LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGIONS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X69198; CAA49137.1; -.
DR EMBL; X67117; CAA47540.1; -.
DR PIR; D36858; D36858.
DR PIR; S35987; S35987.
DR PIR; S46888; S46888.
DR HSP; P19438; INCF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_c6; 2.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00500; TNFR_NGFR_2; 2.
KW Repeat.
FT DOMAIN 31 108 2 x TNFR-CYS.
FT REPEAT 31 66 TNFR-CYS 1.
FT REPEAT 67 108 TNFR-CYS 2.
SQ SEQUENCE 349 AA; 38189 MW; D45D40B5C6E780EF CRC64;

Query Match 36.3%; Score 45; DB 1; Length 349;
Best Local Similarity 36.8%; Pred. No. 15;
Matches 7; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 3 RCRYCNVLCGEREEARACH 21
   ||| | | | | |
Db 83 CLSCNGRCNSQVETRSCN 101

RESULT 10
```

```
PPCK_NEOFR STANDARD; PRT; 608 AA.
ID PCK2130;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE (PHOSPHOENOLPYRUVATE CARBOXYKINASE [GTP] (EC 4.1.1.32)
DE (PHOSPHOENOLPYRUVATE CARBOXYLASE) (PEPCK).
OS Neocallimastix frontalis (Rumen fungus).
OC Eukaryota; Fungi; Chytridiomycota; Neocallimasticales;
OC Neocallimastaceae; Neocallimastix.
CC [1]
CC SEQUENCE FROM N.A.
RX MEDLINE=92184115; PubMed=1339359;
RA Raymond P., Geourjon C., Roux B., Durand R., Fevre M.;
RT "Sequence of the phosphoenolpyruvate carboxykinase-encoding cDNA from
   the rumen anaerobic fungus Neocallimastix frontalis: comparison of
   the amino acid sequence with animals and yeast.";
RL Gene 110:57-63(1992).
CC -1- CATALYTIC ACTIVITY: GTP + OXALOACETATE -> GDP + PHOSPHOENOLPYRUVATE
   + CO(2).
CC -1- PATHWAY: RATE-LIMITING GLUCONEOGENIC ENZYME.
CC -1- SIMILARITY: BELONGS TO THE PHOSPHOENOLPYRUVATE CARBOXYKINASE [GTP]
   FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M59372; AAA33553.1; -.
DR PIR; J01462; J01462.
DR INTERPRO; IPR000364; -.
DR PFAM; PF00821; PEPCK; 1.
DR PROSITE; PS00505; PEPCK_GTP; 1.
KW Gluconeogenesis; Lyase; Decarboxylase; GTP-binding.
FT NP_BIND 217 224 GTP (POTENTIAL).
FT ACT_SITE 269 269 POTENTIAL.
SQ SEQUENCE 608 AA; 66904 MW; C4D2B249A92B7D26 CRC64;

Query Match 35.9%; Score 44.5; DB 1; Length 608;
Best Local Similarity 47.1%; Pred. No. 30;
Matches 8; Conservative 4; Mismatches 2; Indels 3; Gaps 1;

Qy 1 ERCRYCNVLCGEREEA 17
   | | | : | | | | |
Db 77 EKCIY---ICSEKEEDA 90

RESULT 11
FSA_HORSE STANDARD; PRT; 344 AA.
ID FSA_HORSE
AC O62650;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE FOLLISTATIN PRECURSOR (FS) (ACTIVIN-BINDING PROTEIN).
GN FST.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
CC [1]
CC SEQUENCE FROM N.A.
CC TISSUE=FOLLICLE.
RA Sugawara Y., Yamanouchi K.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: BINDS DIRECTLY TO ACTIVIN. SPECIFIC INHIBITOR OF THE
CC BIOSYNTHESIS AND SECRETION OF PITUITARY FOLLICLE STIMULATING
CC HORMONE (FSH).
CC
```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AB010829; BAA25699.1; -
DR INTERPRO: IPR002350; -
DR PFAM: PF00050; kazal.3.
KW Signal; Glycoprotein.
FT SIGNAL 1 29 POTENTIAL.
FT CHAIN 30 344 FOLLISTATIN.
FT CARBOHYD 124 124 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 288 288 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 344 AA; 38002 MW; 1803577D2D9BE4AA CRC64;

Query Match 35.5%; Score 44; DB 1; Length 344;
Best Local Similarity 36.8%; Pred. No. 21;
Matches 7; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 2 RCYCNVLCGEREEARAC 20
Db 266 RCSLCLDELCPDSKSEPPVC 284

RESULT 12
ID YKHI_CAEEL STANDARD; PRT; 484 AA.
AC P34269;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE HYPOTHETICAL TYROSINASE-LIKE PROTEIN C02C2.1 IN CHROMOSOME III.
GN C02C2.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-BRISTOL N2;
RX MEDLINE-94150718; PubMed-7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kersey J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Ropra A., Saunders D., Showkneen R.,
RA Sims M., Smalton N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterson R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans".
RL Nature 368:32-38(1994).
CC -1- COFACTOR: BINDS TWO COPPER IONS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE TYROSINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L23649; AAA27909.1; -
DR PIR: S44739; S44739.
DR WORMPEP; C02C2.1; CE06754.
DR -----

DR INTERPRO: IPR002227; -
DR PFAM: PF00264; tyrosinase.1.
DR PRINTS: PR00092; TYROSINASE.
DR PROSITE: PS00497; TYROSINASE_1; 1.
DR PROSITE: PS00498; TYROSINASE_2; 1.
KW Hypothetical protein; oxidoreductase; Monooxygenase; Copper.
FT METAL 41 41 COPPER A (BY SIMILARITY).
FT METAL 50 50 COPPER A (BY SIMILARITY).
FT METAL 172 172 COPPER B (BY SIMILARITY).
FT METAL 176 176 COPPER B (BY SIMILARITY).
FT METAL 199 199 COPPER B (BY SIMILARITY).
SQ SEQUENCE 484 AA; 55509 MW; 2A9DA4EC86CCF3F6 CRC64;

Query Match 35.5%; Score 44; DB 1; Length 484;
Best Local Similarity 40.0%; Pred. No. 29;
Matches 8; Conservative 2; Mismatches 8; Indels 2; Gaps 1;

QY 4 RYCNVLCG--EREERACH 21
Db 388 RYCRKSCGLCOSNDNNGCH 407

RESULT 13
ID MUC2_HUMAN STANDARD; PRT; 5179 AA.
AC Q02817; Q14878;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE MUCIN 2 PRECURSOR (INTESTINAL MUCIN 2).
GN MUC2 OR SMUC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=INTESTINE;
RX MEDLINE-94132002; PubMed-8300571;
RA Gum J.R. Jr., Hicks J.W., Toribara N.W., Siddiki B., Kim Y.S.;
RT "Molecular cloning of human intestinal mucin (MUC2) cDNA.
RT Identification of the amino terminus and overall sequence similarity
RT to prepro-von Willebrand factor".
RL J. Biol. Chem. 269:2440-2446(1994).
RN [2]
RP SEQUENCE OF 626-1895 AND 4196-5179 FROM N.A.
RC TISSUE=COLON;
RX MEDLINE-93016075; PubMed-1400449;
RA Gum J.R. Jr., Hicks J.W., Toribara N.W., Rothe E.-M., Lagace R.E.,
RA Kim Y.S.;
RT "The human MUC2 intestinal mucin has cysteine-rich subdomains located
RT both upstream and downstream of its central repetitive region".
RL J. Biol. Chem. 267:21375-21383(1992).
RN [3]
RP SEQUENCE OF 1343-1895 AND 4176-4195 FROM N.A.
RX MEDLINE-91358717; PubMed-1885763;
RA Toribara N.W., Gum J.R. Jr., Culhane P.J., Lagace R.E., Hicks J.W.,
RA Petersen G.M., Kim Y.S.;
RT "MUC-2 human small intestinal mucin gene structure. Repeated arrays
RT and polymorphism".
RL J. Clin. Invest. 88:1005-1013(1991).
CC -1- FUNCTION: COATS THE EPITHELIA OF THE INTESTINES, AIRWAYS, AND
CC OTHER MUCUS MEMBRANE-CONTAINING ORGANS. THOUGHT TO PROVIDE A
CC PROTECTIVE, LUBRICATING BARRIER AGAINST PARTICLES AND INFECTIOUS
CC AGENTS AT MUCOSAL SURFACES.
CC -1- SUBUNIT: MULTIMERIC.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- TISSUE SPECIFICITY: COLON, SMALL INTESTINE, COLONIC TUMORS,
CC BRONCHUS, CERVIX AND GALL BLADDER.
CC -1- PTM: ALL CYSTEINE RESIDUES ARE INVOLVED IN INTRACHAIN OR
CC INTERCHAIN DISULFIDE BONDS (BY SIMILARITY).
CC -1- POLYMORPHISM: THE NUMBER OF REPEATS IS HIGHLY POLYMORPHIC AND
CC VARIES AMONG DIFFERENT ALLELES.
CC -----


```
CC      HAS PARALYTIC ACTIVITY IN MICE.
CC      -!- SUBCELLULAR LOCATION: SECRETED.
CC      -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC
CC      PIR; JT0014; NTSR9E.
DR      HSP; P01487; 1LQO.
DR
DR      INTERPRO: IPR002061; -.
DR      PFAM; PF00537; toxin_3; 1.
DR      PRINTS; PR00285; SCORPNTOXIN.
KW      Neurotoxin; Sodium channel inhibitor.
FT      DISULFID 12 65
FT      DISULFID 16 38
FT      DISULFID 24 48
FT      DISULFID 28 50
SQ      SEQUENCE 66 AA; 7343 MW; B8ADB5FE47AE263 CRC64;

Query Match      34.7%; Score 43; DB 1; Length 66;
Best Local Similarity 50.0%; Pred. No. 6.5;
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy      5 YCNVLCGEREEEARAC 20
      ||| || | | |
Db      23 YCNDLCTENGAEAGYC 38

RESULT 15
FIXX_RHILT
ID      FIXX_RHILT      STANDARD;      PRT;      98 AA.
AC      P08710;
DT      01-JAN-1988 (Rel. 06, Created)
DT      01-JAN-1988 (Rel. 06, Last sequence update)
DT      30-MAY-2000 (Rel. 39, Last annotation update)
DE      FERREDOXIN LIKE PROTEIN.
GN      FIXX.
OS      Rhizobium leguminosarum (biovar trifolii).
OG      plasmid sym.
OC      Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC      Rhizobiaceae; Rhizobium.
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ANO 843;
RX      MEDLINE=87174837; PubMed=3562251;
RA      Iismaa S.E., Watson J.M.;
RT      "A gene upstream of the Rhizobium trifolii nifA gene encodes a
      ferredoxin-like protein.";
RL      Nucleic Acids Res. 15:3180-3180(1987).
CC      -!- FUNCTION: COULD BE A 3FE-4S CLUSTER-CONTAINING PROTEIN.
CC      -!- SIMILARITY: TO FERREDOXINS FROM PSEUDOMONAS PUTIDA AND CLOSTRIDIUM
      TARTARIVORUM, FERREDOXIN I FROM ACETOBACTER VINELANDII, FERREDOXIN
      II FROM DESULFOVIBRIO DESULFURICANS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
      between the Swiss Institute of Bioinformatics and the EMBL outstation -
      the European Bioinformatics Institute. There are no restrictions on its
      use by non-profit institutions as long as its content is in no way
      modified and this statement is not removed. Usage by and for commercial
      entities requires a license agreement (See http://www.isb-sib.ch/announce/
      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; X05257; CAA28979.1; -.
DR      PIR; A27510; A27510.
DR      INTERPRO; IPR001450; -.
DR      PROSITE; PS00198; 4FE4S_FERREDOXIN; 1.
KW      Electron transport; Iron-sulfur; Nitrogen fixation; plasmid.
SQ      SEQUENCE 98 AA; 11110 MW; 6BC3047472E1F764 CRC64;

Query Match      34.7%; Score 43; DB 1; Length 98;
Best Local Similarity 77.8%; Pred. No. 9.3;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      3 CRYCNVLCG 11
```

Db 69 CGTCNVLCG 77

Search completed: January 30, 2001, 17:06:11
Job time: 1232 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:55:00 ; Search time 254.27 seconds
(without alignments)
9.680 Million cell updates/sec

Title: US-09-518-931-4_COPY_86_106
Perfect score: 124
Sequence: 1 ECRYCNVLCGEREEARACH 21

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL15:
1: sp-archaea:
2: sp-bacteria:
3: sp-fungi:
4: sp-human:
5: sp-invertebrate:
6: sp-mammal:
7: sp-mhc:
8: sp-organelle:
9: sp-phase:
10: sp-plant:
11: sp-rodent:
12: sp-virus:
13: sp-vertebrate:
14: sp-unclassified:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	124	100.0	300	4	O95407
2	52.5	42.3	268	4	Q14468
3	52.5	42.3	318	4	Q15779
4	51	41.1	319	13	Q91376
5	51	41.1	469	1	Q30094
6	50	40.3	182	4	Q9NT42
7	50	40.3	1696	5	Q9NJ15
8	49	39.5	150	5	Q9VP99
9	49	39.5	1323	5	Q9NJ14
10	49	39.5	1343	5	Q9NJ16
11	48	38.7	202	12	Q9QC04
12	48	38.7	657	2	Q9KRK4
13	47.5	38.3	347	5	Q9VX61
14	47	37.9	114	1	Q58738
15	47	37.9	377	2	Q9S5A3
16	47	37.9	446	5	Q9UAX0
17	46.5	37.5	313	5	Q9XUJ6
18	46	37.1	336	2	Q9WZ28
19	46	37.1	409	5	Q9V756

20	46	37.1	488	5	O44751
21	45.5	36.7	2527	5	O24107
22	45.5	36.7	2529	5	O24605
23	45.5	36.7	2532	5	O9V5T5
24	45	36.3	343	13	Q90844
25	45	36.3	343	13	Q9PS97
26	45	36.3	348	12	O57277
27	45	36.3	348	12	O57103
28	45	36.3	348	12	O57108
29	45	36.3	348	12	O57112
30	45	36.3	348	12	O85407
31	45	36.3	349	12	O57291
32	45	36.3	349	12	O57099
33	45	36.3	349	12	O57100
34	45	36.3	349	12	O57101
35	45	36.3	349	12	O57102
36	45	36.3	349	12	O57109
37	45	36.3	349	12	O57110
38	45	36.3	349	12	O57111
39	45	36.3	349	12	O89118
40	45	36.3	349	12	O89098
41	45	36.3	384	2	P71055
42	45	36.3	418	3	Q9USM3
43	45	36.3	483	5	Q9VUG2
44	45	36.3	586	5	Q9N6P9
45	44.5	35.9	539	11	Q60664

ALIGNMENTS

RESULT 1
O95407
ID O95407 PRELIMINARY; PRT: 300 AA.
AC O95407;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE DECOY RECEPTOR 3 (M68) (M68C) (M68E).
GN DCR3 OR Tr6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE FROM N.A.
RX MEDLINE-99087326; PubMed-9872321;
RA Pitti R.M., Marsters S.A., Lawrence D.A., Roy M., Kischkel F.C.,
RA Dowd P., Huang A., Donahue C.J., Sherwood S.W., Baldwin D.T.,
RA Godowski P.J., Wood W.I., Gurney A.L., Hillan K.J., Cohen R.L.,
RA Goddard A.D., Botstein D., Ashkenazi A.;
RT "Genomic amplification of a decoy receptor for Fas ligand in lung and
RT colon cancer.";
RL Nature 396:699-703(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX SEQUENCE-BLOOD;
RC TISSUE-BLOOD;
RX MEDLINE-99253915; PubMed-10318773;
RA Yu K.Y., Kwon B., Ni J., Zhai Y., Ebner R., Kwon B.S.;
RT "A newly identified member of tumor necrosis factor receptor
RT superfamily (TR6) suppresses LIGHT-mediated apoptosis.";
RL J. Biol. Chem. 274:13733-13736(1999).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE-PANCREAS;
RC MEDLINE-20122600; PubMed-10655513;
RA Bai C., Connolly B., Metzker M.L., Hilliard C.A., Liu X., Sandig V.,
RA Soderman A., Galloway S.M., Liu Q., Austin C.P., Caskey C.T.;
RT "Overexpression of M68/Dcr3 in human gastrointestinal tract tumors
RT independent of gene amplification and its location in a four-gene
RT cluster.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:1230-1235(2000).
DR EMBL; AF104419; AAD03056.1; -

DR	EMBL; AF134240; AAD29688.1; -	OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
DR	EMBL; AF217796; AAF35244.1; -	OX	NCBI_TaxID=9606;
DR	EMBL; AF217793; AAF33685.1; -	RN	[1]
DR	EMBL; AF217794; AAF33686.1; -	RP	SEQUENCE FROM N.A.
DR	HSSP; P25942; 1CDF	RX	MEDLINE=97017139; PubMed=8863740;
DR	INTERPRO; IPR000561; -	RA	Scott I.C., Halli R., Jenkins J.M., Mehan S., Apostolou S.,
DR	PFAM; PF00020; TNFR_C6; 4.	RA	Wingqvist R., Callen D.F., Prockop D.J., Peltonen L., Kadler K.E.;
DR	PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.	RT	"Molecular cloning, expression and chromosomal localization of a human
DR	PROSITE; PS01186; EGF_2; UNKNOWN_1.	RT	gene encoding a 33 kDa putative metalloproteinase (PRSM1).";
DR	PROSITE; PS00050; TNFR_NGFR_2; 2.	RL	Gene 174:135-143(1996).
DR	PRODOM; PD000771; -; 1.	DR	EMBL; U58048; AAC50775.1; -
KW	Receptor.	DR	INTERPRO; IPR000130; -
SQ	SEQUENCE 300 AA; 32679 MW; F90AE33718449AF CRC64;	DR	PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
		SQ	SEQUENCE 318 AA; 33249 MW; A2471BD7293C7574 CRC64;
Query Match	100.0%; Score 124; DB 4; Length 300;	Query Match	42.3%; Score 52.5; DB 4; Length 318;
Best Local Similarity	100.0%; Pred. No. 2,1e-10;	Best Local Similarity	47.6%; Pred. No. 4.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		Matches 10; Conservative 3; Mismatches 7; Indels 1; Gaps 1;	
QY	1 ECRYCNVLCGEREEARACH 21	QY	1 ECRYCNVLCGERE-EEARAC 20
Db	86 ECRYCNVLCGEREEARACH 106	Db	111 EACRVPCVCRERHPQERRC 131
RESULT 2		RESULT 4	
ID	Q14468 PRELIMINARY; PRT; 268 AA.	ID	Q91376 PRELIMINARY; PRT; 319 AA.
AC	Q14468;	AC	Q91376;
DT	01-NOV-1996 (TRENBLrel. 01, Created)	DT	01-NOV-1996 (TRENBLrel. 01, Created)
DT	01-NOV-1996 (TRENBLrel. 01, Last sequence update)	DT	01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT	01-MAY-2000 (TRENBLrel. 13, Last annotation update)	DT	01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE	KIAA0047 PROTEIN (FRAGMENT).	DE	FOLLISTATIN.
GN	KIAA0047.	OS	Xenopus laevis (African clawed frog).
OS	Homo sapiens (Human).	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	OC	Xenopodinae; Xenopus.
OX	NCBI_TaxID=9606;	OX	NCBI_TaxID=8355;
RN	[1]	RN	[1]
RP	SEQUENCE FROM N.A.	RP	SEQUENCE FROM N.A.
RX	MEDLINE=96051398; PubMed=7584044;	RX	MEDLINE=94221645; PubMed=8168135;
RA	Nomura N., Nagase T., Miyajima N., Sazuka T., Tanaka A., Sato S.,	RA	Hemmati-Brivanlou A., Kelly O.G., Melton D.A.;
RA	Seki N., Kawarabayashi Y., Ishikawa K., Tabata S.;	RT	"Follistatin, an antagonist of activin, is expressed in the Spemann
RT	"Prediction of the coding sequences of unidentified human genes. II.	RT	organizer and displays direct neutralizing activity.";
RT	The coding sequences of 40 new genes (KIAA0041-KIAA0080) deduced by	RL	Cell 77:283-295(1994).
RT	analysis of cDNA clones from human cell line KG-1.";	RL	EMBL; S69801; AAB30638.1; -
RL	DNA Res. 1:223-229(1994).	DR	HSSP; P09486; LBMO.
DR	EMBL; D38554; BAA07557.1; -	DR	INTERPRO; IPR002350; -
DR	INTERPRO; IPR000130; -	DR	PFAM; PF00050; Kazal; 3.
DR	PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.	DR	PFAM; PF00050; Kazal; 3.
FT	NON_TER	DR	PFAM; PF00050; Kazal; 3.
FT	NON_TER	DR	PFAM; PF00050; Kazal; 3.
SQ	SEQUENCE 268 AA; 27850 MW; AE187A86CFD0E314 CRC64;	SQ	SEQUENCE 319 AA; 34914 MW; D2CF5E72E30E8EA9 CRC64;
Query Match	42.3%; Score 52.5; DB 4; Length 268;	Query Match	41.1%; Score 51; DB 13; Length 319;
Best Local Similarity	47.6%; Pred. No. 3.5;	Best Local Similarity	42.1%; Pred. No. 6.8;
Matches 10; Conservative 3; Mismatches 7; Indels 1; Gaps 1;		Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;	
QY	1 ECRYCNVLCGERE-EEARAC 20	QY	2 RCRYCNVLCGEREEARAC 20
Db	61 ECRVPCVCRERHPQERRC 81	Db	266 RCALCDLCLGSKSDTVC 284
RESULT 3		RESULT 5	
ID	Q15779 PRELIMINARY; PRT; 318 AA.	ID	Q30094 PRELIMINARY; PRT; 469 AA.
AC	Q15779;	AC	Q30094;
DT	01-NOV-1996 (TRENBLrel. 01, Created)	DT	01-JAN-1998 (TRENBLrel. 05, Created)
DT	01-NOV-1996 (TRENBLrel. 01, Last sequence update)	DT	01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT	01-NOV-1996 (TRENBLrel. 01, Last sequence update)	DT	01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE	PRSM1.	DE	CONSERVED HYPOTHETICAL PROTEIN.
OS	Homo sapiens (Human).	GN	AF0143.
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	OS	Archaeoglobus fulgidus.
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	OC	Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	OC	Archaeoglobus.

```
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=938475;
RA Klenk H.-B., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.T., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
DR EMBL; AE001096; AAB91081.1; -
DR TIGR; AF0143; -
DR INTERPRO; IPR001064; -
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 469 AA; 52627 MW; FD2485B8209D4F52 CRC64;

Query Match 41.1%; Score 51; DB 1; Length 469;
Best Local Similarity 45.0%; Pred. No. 9.5;
Matches 9; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 1 ERCRYCNVLCGEREEARAC 20
   I: I I I I I I I I I I
DB 195 EKRCRCRCEFLKGSEGYAC 214

RESULT 6
Q9NT42 PRELIMINARY; PRT; 182 AA.
AC Q9NT42;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TremBLrel. 15, Last annotation update)
DE HYPOTHETICAL.20.4 KDA PROTEIN (FRAGMENT).
GN DKF2P434A1520
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Ottenwaelder B., Obermaier B., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL137544; CAB70802.1; -
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 182 AA; 20381 MW; 5045A0B3301A991C CRC64;

Query Match 40.3%; Score 50; DB 4; Length 182;
Best Local Similarity 44.4%; Pred. No. 5.7;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 CRYCNVLCGEREEARAC 20
   I: I I I I I I I I I I
DB 165 CKRCCKVMLGETVSSARS 182

RESULT 7
Q9NJ15 PRELIMINARY; PRT; 1696 AA.
ID Q9NJ15
AC Q9NJ15
```

```
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TremBLrel. 15, Last annotation update)
DE PROPROTEIN CONVERTASE APC6B ISOFORM.
GN PC6.
OS Branchiostoma californiensis (California lancelet) (Amphioxus).
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC Branchiostoma.
OX NCBI_TaxID=7738;
RN [1]
RP SEQUENCE FROM N.A.
RA Oliva A.A. Jr., Chan S.-J., Steiner D.F.;
RT "Evolution of the prohormone convertases: identification of a
homologue of PC6 in the protochordate amphioxus.";
RL Biochim. Biophys. Acta 1477:338-348(2000).
DR EMBL; AF184616; AAF26301.1; -
SQ SEQUENCE 1696 AA; 188409 MW; 281CBE1784257CBD CRC64;

Query Match 40.3%; Score 50; DB 5; Length 1696;
Best Local Similarity 43.5%; Pred. No. 41;
Matches 10; Conservative 3; Mismatches 6; Indels 4; Gaps 1;

QY 2 RCRCYNVLC-----GEREEARAC 20
   : I I I I I I I I I I
DB 1388 QCRLCVNCVCKTCHGEDECMEC 1410

RESULT 8
Q9VP99 PRELIMINARY; PRT; 150 AA.
ID Q9VP99
AC Q9VP99;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TremBLrel. 15, Last annotation update)
DE CGI3257 PROTEIN.
GN CGI3257
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Phlebotomidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer V.G., Champetier M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Iredow C.,
RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
```

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Spier B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Shue E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveril J.S., Zhan M., Zhang G., Zhao X., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Zhong H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003592; AAF51656.1; -.
DR HSP: O07954; 1CR8.
DR FLYBASE: FBgn037034; CG13257.
DR INTERPRO: IPR001545; -.
DR INTERPRO: IPR002172; -.
DR PFAM: PF00057; 1dl_recept_a; 1.
DR PROSITE: PS00261; GLYCO_HORMONE_BETA_1; UNKNOWN_1.
DR PROSITE: PS01209; LDLRA_1; 1.
DR PROSITE: PS00668; LDLRA_2; 1.
DR SEQUENCE 150 AA: 15131 MW: 7A065C21C1376079 CRC64:

```
Query Match      39.5%   Score 49; DB 5; Length 150;
Best Local Similarity 44.0%; Pred. No. 6.7;
Matches 11; Conservative 3; Mismatches 5; Indels 6; Gaps 2;
```

```

RESULT      9
Q9NJ14      PRELIMINARY;      PRT;      1323 AA.
AC          Q9NJ14;
DT          01-OCT-2000 (TrEMBLrel. 15, Created)
DT          01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT          01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE          DE. PROTEIN CONVERTASE APC6C ISOFORM.
PC          PC6.
OS          Branchiostoma californiensis (California lancelet) (Amphioxus).
OC          Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC          Branchiostoma.
NCBI_TaxID=7738;
[1]
RN          SEQUENCE FROM N A.
RP          OLIVA A. A. Jr., Chan S. J., Steiner D. F.;
RA          "Evolution of the prohormone convertases: identification of a
RT          homologue of PC6 in the protochordate amphioxus.";
RL          Biochim. Biophys. Acta 1477:338-348(2000).
DR          EMBL; AF184617; AAF26302.1;
SQ          SEQUENCE 1323 AA; 146039 MW; 75222694B06951D CRC64;

```

Query Match	39.5%	Score 49;	DB 5;	Length 1323;
Best Local Similarity	29.2%	Pred. No. 46;		
Matches	7;	Conservative	6;	Mismatches
			7;	Indels
				Gaps

```

Qy      2  RCRYCNVLQ----GEREEERACH  21
      :|  ||::|  |  :  :  ||
Db      873  KCEDCNIIICKNGSQADQCLECH  896

RESULT  10
Q9NJ16
ID      Q9NJ16      PRELIMINARY;      PRT;  1343  AA.
AC      Q9NJ16;
DT      01-OCT-2000  (TREMBLrel. 15, Created)
DT      01-OCT-2000  (TREMBLrel. 15, Last sequence update)
DT      01-OCT-2000  (TREMBLrel. 15, Last annotation update)
DE      PROTEIN CONVERTASE APC6A ISOFORM.
GN      PC6.

```

OS Branchiostoma californiensis (California lancelet) (Amphioxus).
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC Branchiostoma.
ON NCBI_TaxID=7738;
RX [1]
RP SEQUENCE FROM N.A.
RA Oliva A.A. Jr., Chan S.J., Steiner D.F.;
RT "Evolution of the prohomone convertases: identification of a
RT homologue of pc6 in the protochordate amphioxus.";
RL Biochim. Biophys. Acta 1477:338-348(2000).
DR EMBL; AF184615; AAF26300.1; -
SQ SEQUENCE 1343 AA; 149307 MW; 1359D68401FA07FE CRC64;

```

Query Match      39.5%; Score 49; DB 5; Length 1343;
Best Local Similarity 29.2%; Pred. No. 46;
Matches 7; Conservative 6; Mismatches 7; Indels

Qy  2  RCRCYNVLC-----GREGEEARACH  21
      :| |||:| :| :|
Db  873  KCEDCNIICKKNGSOADOCLECH  896

```

RESULT	11
Q9QC04	
ID	O9QC04 PRELIMINARY; PRT; 202 AA.
AC	O9QC04;
DT	01-MAY-2000 (TrEMBLrel. 13, Created)
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT	01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE	DE SPLF4 PROTEIN.
GN	SPLF4.
OS	Simian cytomegalovirus (strain Colburn).
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC	Betaherpesvirinae; Cytomegalovirus.
OX	NCBI_TaxID=50292;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	STRATN=COLBURN;
RC	Zong J.-C., Jeang K.T., Hayward G.S.;
RT	"Sequence of the Sali-P fragment of simian cytomegalovirus strain
RT	Colburn containing CA-repeats.";
RL	Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR	EMBL; J02062; AAF08542.1; --
DZ	SEQUENCE 202 AA; 22786 MW; 7D611193F959E58D3 CRC64;
SQ	

Query Match	Best Local Similarity	Score	DB	Length
38.7%	36.8%	48	12	202
Matches	7	Conservative	4	Mismatches
			8	Indels
			0	Gaps

[illegible]


```
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Raogoi I., Sellers P.,
RA McDonald L., Uterback T., Fleishmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
cholerae.";
RL Nature 406:477-483(2000).
DR EMBL; AE004241; AAF94787.1; -.
DR TIGR; VC1636; -.
SQ SEQUENCE 657 AA; 74226 MW; 57FC7D16F01D7D71 CRC64;

Query Match 38.7%; Score 48; DB 2; Length 657;
Best Local Similarity 40.0%; Pred. No. 34;
Matches 8; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 2 RCRCYNVLCGEREEARACH 21
I :||| : : |||
Db 493 RAKYCNCGADNDIAARICH 512

RESULT 13
Q9VX61
ID Q9VX61 PRELIMINARY; PRT; 347 AA.
AC Q9VX61
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE CG12996 PROTEIN.
GN CG12996.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foshler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kimms I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
```

```
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003505; AAF48717.1; -.
DR FLYBASE; FBgn0030835; CG12996.
SQ SEQUENCE 347 AA; 38803 MW; 681F197A01FA0D3C CRC64;

Query Match 38.3%; Score 47.5; DB 5; Length 347;
Best Local Similarity 45.0%; Pred. No. 23;
Matches 9; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

QY 3 CRYC-NVLCGEREEARACH 21
I :||| : |||
Db 257 CRFCENSLCGSKCNSRGSH 276

RESULT 14
Q58738
ID Q58738 PRELIMINARY; PRT; 114 AA.
AC Q58738;
DT 01-JUN-1998 (TRENBLrel. 06, Created)
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
DT 01-JUN-1998 (TRENBLrel. 06, Last annotation update)
DE HYPOTHETICAL PROTEIN MJ1342.
GN MJ1342.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 2661;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii.";
RL Science 273:1058-1073(1996).
CC -!- SIMILARITY: WEAK, TO M.JANNASCHII MJ0207 AND MJ1460.
DR EMBL; U67574; AAB99359.1; -.
DR TIGR; MJ1342; -.
KW Hypothetical protein.
SQ SEQUENCE 114 AA; 13134 MW; 73DB2C0F38AA616A CRC64;

Query Match 37.9%; Score 47; DB 1; Length 114;
Best Local Similarity 40.0%; Pred. No. 10;
Matches 8; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 ERCRCYNVLCGEREEARAC 20
I :||| : |||
Db 54 DKCKYCPKQIGAEIPGSAC 73

RESULT 15
Q9S5A3
ID Q9S5A3 PRELIMINARY; PRT; 377 AA.
AC Q9S5A3;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE DNAJ.
GN DNAJ.
```

OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=10403S;
RA Hanawa T., Kai M., Kamiya S., Yamamoto T.;
RT "Cloning, sequencing, and transcriptional analysis of the dnaK heat
RT shock operon of Listeria monocytogenes.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB023064; BAA82790.1; -
DR HSSP; P08622; LBQZ.
DR INTERPRO: IPR001305; -
DR INTERPRO: IPR001623; -
DR INTERPRO: IPR002939; -
DR INTERPRO: IPR003095; -
DR PFAM; PF00226; DnaJ; 1.
DR PFAM; PF00684; DnaJ_CXXCXGXG; 1.
DR PFAM; PF01556; DnaJ_C; 1.
DR PRINTS; PR00625; DnaJPROTEIN.
DR PROSITE; PS00636; DnaJ_1; 1.
DR PROSITE; PS00637; DnaJ_CXXCXGXG; 1.
DR PROSITE; PS0076; DnaJ_2; 1.
SQ SEQUENCE 377 AA; 41077 MW; 70A8C5836239310A CRC64;

Query Match 37.9%; Score 47; DB 2; Length 377;
Best Local Similarity 42.1%; Pred. No. 29;
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
QY 3 CRYCNVLCGEREEERACH 21
|:|:| | | | | | | | | |
Db 190 CQVCNGTGKEIKEKCPICH 208

Search completed: January 30, 2001, 16:55:02
Job time: 918 sec

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ERRCYCNVLCGEREEARACH 21
|||||
Db 86 ERRCYCNVLCGEREEARACH 106

RESULT 2
US-08-152-019A-26
; Sequence 26, Application US/08152019A
; Patent No. 5565331
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: Serafini, Tito
; APPLICANT: Kennedy, Timothy
; APPLICANT: Placzek, Marysia
; APPLICANT: Jessel, Thomas
; APPLICANT: Dodd, Jane
; TITLE OF INVENTION: NEURAL AXON OUTGROWTH MODULATORS
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/152,019A
; FILING DATE: 12-NOV-1993
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard Aron
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-59012/RAO
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299 FHT UR
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 605 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-152-019A-26

Query Match 37.5%; Score 46.5; DB 1; Length 605;
Best Local Similarity 50.0%; Pred. No. 45;
Matches 9; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 4 RYCYNVLCGEREEARACH 21
|||||
Db 77 RYC-VVTEKGEEQVRSCH 93

RESULT 3
US-08-482-677-4
; Sequence 4, Application US/08482677
; Patent No. 6017714
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: Serafini, Tito
; APPLICANT: Kennedy, Timothy
; APPLICANT: Placzek, Marysia
; APPLICANT: Jessel, Thomas

; APPLICANT: Dodd, Jane
; TITLE OF INVENTION: Netrins
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,677
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC93-300-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 605 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-482-677-4

Query Match 37.5%; Score 46.5; DB 3; Length 605;
Best Local Similarity 50.0%; Pred. No. 45;
Matches 9; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 4 RYCYNVLCGEREEARACH 21
|||||
Db 77 RYC-VVTEKGEEQVRSCH 93

RESULT 4
US-08-650-599A-1
; Sequence 1, Application US/08650599A
; Patent No. 6096866
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: Serafini, Tito
; APPLICANT: Kennedy, Timothy
; APPLICANT: Placzek, Marysia
; APPLICANT: Jessel, Thomas
; APPLICANT: Dodd, Jane
; TITLE OF INVENTION: Netrins
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/650,599A
; FILING DATE: 07-JUN-1995

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC93-300-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 605 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-650-599A-1

Query Match 37.5%; Score 46.5; DB 3; Length 605;
Best Local Similarity 50.0%; Pred. No. 45;
Matches 9; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

Qy 4 RYCNVLCGEREEARACH 21
||| | : ||: |:
Db 77 RYC-VVTEKGEEQVRSCH 93

RESULT 5
US-08-665-259-22
Sequence 22, Application US/08665259
Patent No. 6028173
GENERAL INFORMATION:
APPLICANT: Landes, Gregory M.
APPLICANT: Burn, Timothy C.
APPLICANT: Connors, Timothy D.
APPLICANT: Dackowski, William R.
APPLICANT: Van Raay, Terence J.
APPLICANT: Klinger, Katherine W.
TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
METHODS OF MAKING AND USING SAME
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENZYME CORPORATION
STREET: One Mountain Road
CITY: Framingham
STATE: Massachusetts
COUNTRY: United States of America
ZIP: 01701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 17-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dugan, Deborah A.
REGISTRATION NUMBER: 37,315
REFERENCE/DOCKET NUMBER: IG5-9.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508) 872-8400
TELEFAX: (508) 872-5415
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 606 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-665-259-22

Query Match 37.5%; Score 46.5; DB 3; Length 606;
Best Local Similarity 50.0%; Pred. No. 45;
Matches 9; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

Qy 4 RYCNVLCGEREEARACH 21
||| | : ||: |:
Db 78 RYC-VVTEKGEEQVRSCH 94

RESULT 6
US-08-762-500-22
Sequence 22, Application US/08762500
Patent No. 6030806
GENERAL INFORMATION:
APPLICANT: Landes, Gregory M.
APPLICANT: Burn, Timothy C.
APPLICANT: Connors, Timothy D.
APPLICANT: Dackowski, William R.
APPLICANT: Van Raay, Terence J.
APPLICANT: Klinger, Katherine W.
TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
METHODS OF MAKING AND USING SAME
NUMBER OF SEQUENCES: 83
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENZYME CORPORATION
STREET: One Mountain Road
CITY: Framingham
STATE: Massachusetts
COUNTRY: United States of America
ZIP: 01701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 09-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/762,500
FILING DATE: 17-JUN-1996
APPLICATION NUMBER: US 08/665,259
FILING DATE: 17-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/10469
FILING DATE: 17-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Dugan, Deborah A.
REGISTRATION NUMBER: 37,315
REFERENCE/DOCKET NUMBER: IG5-9.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508) 872-8400
TELEFAX: (508) 872-5415
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 606 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-762-500-22

Query Match 37.5%; Score 46.5; DB 3; Length 606;
Best Local Similarity 50.0%; Pred. No. 45;
Matches 9; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

Qy 4 RYCNVLCGEREEARACH 21
||| | : ||: |:
Db 78 RYC-VVTEKGEEQVRSCH 94

RESULT 7

Wed Jan 31 13:45:16 2001

```
US-08-612-986-7
; Sequence 7, Application US/08612986
; Patent No. 5770384
; GENERAL INFORMATION:
; APPLICANT: Elliot J. Androphy
; APPLICANT: Dave E. Breiding
; TITLE OF INVENTION: E2 BINDING PROTEINS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahive & Cockfield
; STREET: 60 State Street, suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,986
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/361,806
; FILING DATE: 22 DEC 1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: NEP-004DV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-612-986-7

Query Match          36.3%; Score 45; DB 1; Length 122;
Best Local Similarity 63.6%; Pred. No. 16;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 11 GEREERARACH 21
    |||||: ||
Db 81 GEREKEKKVCH 91

RESULT 9
PCT-US95-16806A-7
; Sequence 7, Application PC/TUS9516806A
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: E2 Binding Proteins
; NUMBER OF SEQUENCES: 21
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/16806A
; FILING DATE: December 22, 1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/361,806
; FILING DATE: 22-DEC-1994
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-16806A-7

Query Match          36.3%; Score 45; DB 4; Length 122;
Best Local Similarity 63.6%; Pred. No. 16;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 11 GEREERARACH 21
    |||||: ||
Db 81 GEREKEKKVCH 91

RESULT 10
US-08-292-549-6
; Sequence 6, Application US/08292549
; Patent No. 5464938
; GENERAL INFORMATION:
; APPLICANT: Smith, Craig A.
```

; APPLICANT: Goodwin, Raymond G.
; TITLE OF INVENTION: Isolated Viral Protein TNF Antagonists
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/292,549
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/963,330
; FILING DATE: 10/19/92
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia A.
; REGISTRATION NUMBER: 34,693
; REFERENCE/DOCKET NUMBER: 2602-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-292-549-6

Query Match 35.5%; Score 44; DB 1; Length 355;
Best Local Similarity 36.8%; Pred. No. 59;
Matches 7; Conservative 3; Mismatches 9; Indels 0; Gaps 0;
QY 3 CRYCNVLCGEREEERACH 21
| | | | | : | | | |
Db 83 CLSCNGRCDSNQVETRSCN 101

RESULT 11
US-08-477-451-11
; Sequence 11, Application US/08477451
; Patent No. 5928865
; GENERAL INFORMATION:
; APPLICANT: Covacci, Antonello
; TITLE OF INVENTION: Helicobacter Pylori Cagi Region
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,451
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.

; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0335.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2708
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1781 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-477-451-11

Query Match 35.5%; Score 44; DB 2; Length 1781;
Best Local Similarity 75.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 3 CRYCNVLC 10
| | | | | : | |
Db 641 CRYCRLLC 648

RESULT 12
5223423-3
; Patent No. 5223423
; APPLICANT: FRANCHINI, GENOVEFFA;WONG-STAAAL, FLOSSIE;
; GALLO, ROBERT
; TITLE OF INVENTION: CHARACTERIZATION OF REPLICATION COMPETENT
; HUMAN IMMUNODEFICIENCY TYPE 2 PROVIRAL CLONE HIV-2 SBL/ISY
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/331,212
; FILING DATE: 03-31-1989
; SEQ ID NO:3:
; LENGTH: 3025
; 5223423-3

Query Match 35.5%; Score 44; DB 5; Length 3025;
Best Local Similarity 44.4%; Pred. No. 4.6e+02;
Matches 8; Conservative 4; Mismatches 2; Indels 4; Gaps 1;
QY 7 NVLCGER----EEEARAC 20
| : | | | | : : | | |
Db 205 NILCGQRNWNWTDQDQAC 222

RESULT 13
US-08-482-677-8
; Sequence 8, Application US/08482677
; Patent No. 6017714
; GENERAL INFORMATION:
; APPLICANT: Tessier-Lavigne, Marc
; APPLICANT: Serafini, Tito
; APPLICANT: Kennedy, Timothy
; APPLICANT: Placzek, Marysia
; APPLICANT: Jessel, Thomas
; APPLICANT: Dodd, Jane
; TITLE OF INVENTION: Netrins
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

Wed Jan 31 13:45:16 2001

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/482.677

FILING DATE: 07-JUN-1995

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: OSMAN, RICHARD A

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: UC93-300-4

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 343-4341

TELEFAX: (415) 343-4342

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 603 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-482-677-8

Query Match 35.1%; Score 43.5; DB 3; Length 603;

Best Local Similarity 50.0%; Pred. No. 1.1e+02;

Matches 9; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 4 RYCNVLCGEREEERACH 21

||| | : |||

DB 75 RYC-VVSGEERVRSC 91

RESULT 14

US-08-459-568-62

; Sequence 62, Application US/08459568

; Patent No. 5811304

; GENERAL INFORMATION:

; APPLICANT: Huang, Shi

; TITLE OF INVENTION: Retinoblastoma Protein - Interacting

; TITLE OF INVENTION: Zinc Finger Proteins

; NUMBER OF SEQUENCES: 93

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Campbell and Flores

; STREET: 4370 La Jolla Village Drive, Suite 700

; CITY: San Diego

; STATE: California

; COUNTRY: USA

; ZIP: 92122

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/459,568

; FILING DATE: 02-JUN-1995

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/399,411

; FILING DATE: 06-MAR-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Campbell, Cathryn A.

; REGISTRATION NUMBER: 31,815

; REFERENCE/DOCKET NUMBER: P-LJ 1264

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619) 535-9001

; TELEFAX: (619) 535-8949

; INFORMATION FOR SEQ ID NO: 62:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 84 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; US-08-459-568-62

Query Match 34.7%; Score 43; DB 2; Length 84;

Best Local Similarity 26.5%; Pred. No. 20;

Matches 9; Conservative 4; Mismatches 7; Indels 14; Gaps 1;

QY 2 RCRYN-----VLCGEREEERACH 21

:|: || | ||: | : ||

DB 28 KCQTCNKGFTQLAHLOKHLYLVHTGKRPHECOVCH 61

RESULT 15

US-08-399-411-62

; Sequence 62, Application US/08399411

; Patent No. 5831008

; GENERAL INFORMATION:

; APPLICANT: Huang, Shi

; TITLE OF INVENTION: Retinoblastoma Protein - Interacting

; TITLE OF INVENTION: Zinc Finger Proteins

; NUMBER OF SEQUENCES: 93

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Campbell and Flores

; STREET: 4370 La Jolla Village Drive, Suite 700

; CITY: San Diego

; STATE: California

; COUNTRY: USA

; ZIP: 92122

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/399,411

; FILING DATE: 06-MAR-1995

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: Campbell, Cathryn A.

; REGISTRATION NUMBER: 31,815

; REFERENCE/DOCKET NUMBER: P-LJ 1264

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619) 535-9001

; TELEFAX: (619) 535-8949

; INFORMATION FOR SEQ ID NO: 62:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 84 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; US-08-399-411-62

Query Match 34.7%; Score 43; DB 2; Length 84;

Best Local Similarity 26.5%; Pred. No. 20;

Matches 9; Conservative 4; Mismatches 7; Indels 14; Gaps 1;

QY 2 RCRYN-----VLCGEREEERACH 21

:|: || | ||: | : ||

DB 28 KCQTCNKGFTQLAHLOKHLYLVHTGKRPHECOVCH 61

Search completed: January 30, 2001, 16:47:50

Job time: 626 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:45:30 ; Search time 183.67 Seconds
(without alignments)
2.234 Million cell updates/sec

Title: US-09-518-931-4_COPY_108_119

Perfect score: 73

Sequence: 1 THNRACRKTGF 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_36:*

- 1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT:*
- 2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT:*
- 3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT:*
- 4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT:*
- 5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT:*
- 6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT:*
- 7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT:*
- 8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT:*
- 9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT:*
- 10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT:*
- 11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT:*
- 12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT:*
- 13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT:*
- 14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT:*
- 15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT:*
- 16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT:*
- 17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT:*
- 18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT:*
- 19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT:*
- 20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT:*
- 21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	73	100.0	153	Y22222	Human TNFR superfa
2	73	100.0	170	W63623	Human tumour necro
3	73	100.0	215	W93585	Human IAP06 protei
4	73	100.0	245	Y28449	A human tumour nec
5	73	100.0	271	Y42184	Human mFLINT #1 pr
6	73	100.0	300	W66102	Amino acid sequenc
7	73	100.0	300	W63622	Human tumour necro
8	73	100.0	300	Y03099	Human lung TNF-rec
9	73	100.0	300	Y42182	Human FLINT #1 pro
10	73	100.0	300	Y17479	Mammalian tumour n
11	73	100.0	300	Y06817	Human DCR3 polypep
12	73	100.0	300	W97749	Human tumour necro

13	73	100.0	300	20	W95082	Orphan receptor (H
14	73	100.0	300	21	Y77458	Human TNF receptor
15	62	84.9	273	20	Y42185	Human mFLINT #2 pr
16	62	84.9	302	20	Y42183	Human FLINT #2 pro
17	56	76.7	360	17	R99938	Mutated OCIF, OCIF
18	49	67.1	143	17	R99946	Mutated OCIF, OCIF
19	49	67.1	145	17	R99930	Osteoclastogenesis
20	49	67.1	154	17	R99929	Osteoclastogenesis
21	49	67.1	187	17	R99950	Mutated OCIF, OCIF
22	49	67.1	187	21	Y77464	Primate protein se
23	49	67.1	197	17	R99945	Mutated OCIF, OCIF
24	49	67.1	272	17	R99944	Mutated OCIF, OCIF
25	49	67.1	321	17	R99949	Mutated OCIF, OCIF
26	49	67.1	326	17	R99940	Mutated OCIF, OCIF
27	49	67.1	327	17	R99941	Mutated OCIF, OCIF
28	49	67.1	349	20	W83928	Human FTHMA-070 pa
29	49	67.1	351	17	R99943	Mutated OCIF, OCIF
30	49	67.1	359	17	R99939	Mutated OCIF, OCIF
31	49	67.1	360	17	R99936	Mutated OCIF, OCIF
32	49	67.1	380	17	R99924	Mature osteoclasto
33	49	67.1	390	17	R99357	Human tumour necro
34	49	67.1	391	19	W52338	Human OCIF genome
35	49	67.1	393	17	R99948	Mutated OCIF, OCIF
36	49	67.1	395	19	W57636	Modified TRI recep
37	49	67.1	399	17	R99942	Mutated OCIF, OCIF
38	49	67.1	401	17	R99932	Mutated OCIF, OCIF
39	49	67.1	401	17	R99933	Mutated OCIF, OCIF
40	49	67.1	401	17	R99934	Mutated OCIF, OCIF
41	49	67.1	401	17	R99935	Mutated OCIF, OCIF
42	49	67.1	401	17	R99925	Full length osteoc
43	49	67.1	401	17	R99931	Mutated OCIF, OCIF
44	49	67.1	401	18	W38345	Human osteoprotege
45	49	67.1	401	20	W83926	Human FTHMA-070 pr

ALIGNMENTS

RESULT 1

Y22222
ID Y22222 standard; Protein; 153 AA.

XX Y22222;

DT 16-SEP-1999 (first entry)

XX Human TNFR superfamily soluble receptor protein sequence.

DE TNFL1; human; TNFR superfamily; tumour necrosis factor ligand; TNF;
KW tumour necrosis factor receptor; TNFR superfamily; cell proliferation;
KW cell differentiation; cytokine production; immunoglobulin; hyperplasia;
KW apoptosis inducer; activated T cell; autoimmune disease; inhibitor;
KW myasthenia gravis; insulin-dependent diabetes mellitus; endotoxin shock;
KW rheumatoid arthritis; multiple sclerosis; systemic lupus erythematosus;
KW tumour; proliferative disorder; neoplasia; dysplasia; immunocompetence;
KW lymphoid organogenesis; bacterial resistance; contact hypersensitivity;
KW delayed type sensitivity; therapy.

XX Homo sapiens.

OS Homo sapiens.

XX WO99933980-A2.

XX 08-JUL-1999.

XX 22-DEC-1998; 98WO-US27474.

XX 16-DEC-1998; 98US-0212270.

XX 30-DEC-1997; 97US-0068959.

XX (CHIR) CHIRON CORP.

XX Kassam A, Lamson G, Pot D, Tribouley C;

Wed Jan 31 13:44:48 2001

```

DR WPI; 1999-405508/34.
DR N-PSDB; X84621.
XX
XX New tumour necrosis factor ligands, useful for induction of cell
XX death and/or proliferation of cells
XX
XX Claim 1; Page 61; 69pp; English.
XX
XX This sequence represents a tumour necrosis factor receptor (TNFR)
XX superfamily soluble protein of the invention. The invention also relates
XX to tumour necrosis factor (TNF) ligand (TNFL) family proteins. The TNFL
XX proteins play regulatory roles in cell proliferation and/or
XX immunoglobulins, etc. A variety of diseases can be treated by modulating
XX the activity of TNFL proteins, e.g. they can induce apoptosis of
XX differentiated, e.g. they can induce production of cytokines,
XX activated T cells but rescue resting T cell from apoptosis. TNFL
XX polypeptides can therefore be used to treat autoimmune diseases, such as
XX myasthenia gravis, insulin-dependent diabetes mellitus, rheumatoid
XX arthritis, multiple sclerosis, and systemic lupus erythematosus. TNFL
XX proteins also have tumour stimulating properties, so tumours can be
XX treated by inhibiting the expression or activity of TNFL. Other
XX proliferative disorders, such as neoplasias, dysplasias, and hyperplasia
XX can also be treated using TNFL inhibitors. The TNFL polypeptides and
XX polynucleotides can also be used to enhance or decrease TNF activity,
XX thus providing therapeutic benefits such as induction of cell death,
XX lymphoid organogenesis, or host bacterial resistance, and inhibition of
XX endotoxin shock, contact hypersensitivity, delayed type sensitivity or
XX immunocompetence of a transplant recipient. TNF and its receptors play a
XX major role in host defence and immunosurveillance. As such, there is a
XX need to identify new members of TNFR families. This invention provides
XX this need.
XX
XX Sequence 153 AA:
XX
XX Query Match 100.0%; Score 73; DB 20; Length 153;
XX Best Local Similarity 100.0%; Pred. No. 0.0011;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 THNRACRCRTGTF 12
XX | | | | | | | | | |
XX Db 24 thnracrcrtg 35
XX
XX RESULT 2
XX W63623 ID W63623 standard; Protein: 170 AA.
XX AC W63623;
XX
XX DT 26-OCT-1998 (first entry)
XX
XX DE Human tumour necrosis factor receptor-6 beta protein.
XX
XX KW Human tumour necrosis factor receptor-6 alpha; TNFR-6 beta;
XX endothelial cells; keratinocytes; normal prostate; apoptosis;
XX prostate tumour tissue.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX FT Peptide 1..30
XX FT Protein 31..170
XX FT note= "TNFR-6 beta"
XX FT Region 31..166
XX FT /note= "Soluble extracellular domain"
XX
XX PN W09830694-A2.
XX
XX PD 16-JUL-1998.
XX
XX PF 13-JAN-1998; 98WO-US00153.
XX
XX

```

```

PR 14-JAN-1997; 97US-0035496.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ebner R, Feng P, Gentz RL, Ni J, Ruben SM, Yu G;
XX WPI; 1998-399142/34.
XX N-PSDB; V39086.
XX
XX Human tumour necrosis factor receptors 6-alpha and 6-beta - used in
XX the diagnosis of immune system-related disorder(s)
XX
XX Claim 20; Fig 2A-2B; 91pp; English.
XX
XX The present sequence represents the human tumour necrosis factor
XX receptor-6 beta (TNFR-6 beta) protein. The invention also provides
XX for the TNFR-6 alpha protein (W63622). TNFR-6 alpha and TNFR-6 beta are
XX members of the tumour necrosis factor receptor (TNFR) family. TNFRs
XX are expressed in endothelial cells, keratinocytes, normal prostate and
XX prostate tumour tissue. For a number of disorders of these cells,
XX particularly of the immune system, substantially altered (whether
XX increased or decreased) levels of TNFR-6 alpha and/or TNFR-6 beta gene
XX expression can be detected, therefore the TNFR-6 alpha and TNFR-6 beta
XX polypeptides, nucleic acids and antibodies are claimed to be useful in
XX the diagnosis of such disorders. Mutations of the TNFR-6 alpha and
XX TNFR-6 beta genes can also be detected. The TNFR polypeptides are
XX also claimed to be useful for identifying ligands which may be useful
XX in the treatment of apoptosis related disorders.
XX
XX Sequence 170 AA:
XX
XX Query Match 100.0%; Score 73; DB 19; Length 170;
XX Best Local Similarity 100.0%; Pred. No. 0.0012;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 THNRACRCRTGTF 12
XX | | | | | | | | | |
XX Db 108 thnracrcrtg 119
XX
XX RESULT 3
XX W93585 ID W93585 standard; Protein: 215 AA.
XX AC W93585;
XX
XX DT 18-JUN-1999 (first entry)
XX
XX DE Human hAPO6 protein.
XX
XX KW Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
XX developmental abnormality; gestational abnormality; prostate cancer;
XX APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
XX cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
XX apoptosis; human.
XX
XX OS Homo sapiens.
XX
XX PN W09911791-A2.
XX
XX PD 11-MAR-1999.
XX
XX PF 04-SEP-1998; 98WO-US18393.
XX
XX PR 05-SEP-1997; 97US-0924634.
XX
XX (UNIW ) UNIV WASHINGTON.
XX
XX Chaudhary PM;
XX WPI; 1999-205191/17.
XX N-PSDB; X23419.
XX

```

XX New Tumor Necrosis Factor family receptor polypeptides and ligands -
 PT useful for diagnosis and treatment of prostate cancer and
 PT developmental or gestational abnormalities
 XX
 XX Claim 29; Fig 9; 156pp; English.
 XX
 XX This invention describes isolated Tumor Necrosis Factor (TNF) family
 CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
 CC their active fragments. APO4 is useful for diagnosing prostate cancer
 CC by determining levels of APO4 in an individual. Prostate cancer can also
 CC be treated using APO4 selective binding agents linked to a therapeutic
 CC moiety. APO4 polypeptides are also useful for identifying selective
 CC binding agents, useful in diagnosis/treatment of disease by binding of
 CC agents to the polypeptide/active fragment which is extracellular, or
 CC expressed on the cell surface. The binding is preferably performed in
 CC vivo. APO4 polypeptides/ active fragments are also useful for screening
 CC for agonists and antagonists by binding and observing the change in APO4
 CC activity. Effective pharmacological agents useful in diagnosis or
 CC treatment of disease are also identified using APO4 polypeptides/active
 CC fragments and APO4 signal transducer molecules that specifically interact
 CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4
 CC activity. The method is performed in vivo or in vitro. APO polypeptides
 CC are all useful as immunogens for preparing antibodies. APO4 is also
 CC useful for diagnosis/treatment of developmental or gestational
 CC abnormalities. APO8 was transfected to human breast carcinoma cell line
 CC MCF-7, and induced apoptosis.
 XX
 XX Sequence 215 AA;

Query Match 100.0%; Score 73; DB 20; Length 215;
 Best Local Similarity 100.0%; Pred. No. 0.0015;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 THNRACRRTGF 12
 |||||
 Db 23 thnracrtgfg 34

RESULT 4
 Y28449
 ID Y28449 standard; Protein; 245 AA.
 XX
 AC Y28449;

XX 29-SEP-1999 (first entry)

XX A human tumour necrosis factor-R2-like proteins (TR2P)-1.

XX Human tumour necrosis factor-R2-like protein; TR2P; achondroplasia;
 KW osteoporosis; developmental disorder; Cushing's syndrome;
 KW muscular dystrophy; epilepsy; hereditary neuropathy;
 KW Charcot-Marie-Tooth disease; neurofibromatosis; hypothyroidism;
 KW hydrocephalus; seizure disorder; cerebral palsy; spinal bifida;
 KW congenital glaucoma; cataract; sensorineural hearing loss;
 KW reproductive disorder; infertility; ovulatory defect; endometriosis;
 KW autoimmune disorder; ectopic pregnancy; teratogenesis; spermatogenesis;
 KW immunological disorder; AIDS; Addison's disease; allergy; bronchitis;
 KW atherosclerosis; diabetes mellitus; Chron's disease; lupus;
 KW irritable bowel syndrome; multiple sclerosis; infection;
 KW neoplastic disorder; adenocarcinoma; leukaemia; lymphoma; melanoma;
 KW myeloma; sarcoma.

XX Homo sapiens.

XX OS WO9931128-A2.

XX PN 24-JUN-1999.

XX PD 02-DEC-1998; 98WO-US25649.

PR 16-DEC-1997; 97US-0991945.
 XX
 XX (INCY-) INCYTE PHARM INC.
 XX

PI Au-Young J, Bandman O, Hillman JL, Kaser MR, Tang YT;
 XX N-PSDB: X89503.

DR WPI; 1999-457916/38.

XX New tumour necrosis factor-R2-like protein - useful in the treatment
 PT of osteogenesis, developmental, reproductive, immunological and
 PT neoplastic disorders

XX Claim 1; Fig 1A-C; 81pp; English.

XX The present sequence represents a human tumour necrosis factor-R2-like
 CC protein (TR2P). The protein is used to treat and prevent osteogenesis,
 CC developmental, reproductive, immunological and neoplastic disorders, and
 CC also to diagnose osteogenesis disorders associated with TR2 protein expression. Such
 CC disorders include osteogenesis disorders such as achondroplasia and
 CC osteoporosis, developmental disorders such as Cushing's syndrome,
 CC muscular dystrophy, and epilepsy, hereditary neuropathies such as
 CC Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism,
 CC hydrocephalus, seizure disorders such as cerebral palsy and spinal
 CC bifida, congenital glaucoma, cataract, or sensorineural hearing loss,
 CC reproductive disorders such as infertility, ovulatory defects and
 CC endometriosis, autoimmune disorders, ectopic pregnancy and teratogenesis,
 CC disruption of spermatogenesis, immunological disorders such as AIDS,
 CC Addison's disease, allergies, bronchitis, atherosclerosis, diabetes
 CC mellitus, Chron's disease, lupus and irritable bowel syndrome, multiple
 CC sclerosis, viral, fungal, helminthic, parasitic and protozoal infections,
 CC and neoplastic disorders including adenocarcinoma, leukaemia, lymphoma,
 CC melanoma, myeloma, sarcoma, and teratocarcinoma.

XX Sequence 245 AA;

Query Match 100.0%; Score 73; DB 20; Length 245;
 Best Local Similarity 100.0%; Pred. No. 0.0017;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 THNRACRRTGF 12
 |||||
 Db 108 thnracrtgfg 119

RESULT 5

Y42184

ID Y42184 standard; Protein; 271 AA.

XX Y42184;

XX 17-DEC-1999 (first entry)

XX Human mFLINT #1 protein sequence.

XX Human; FLINT: mFLINT; OPG3; tumour necrosis factor receptor; FasL;
 KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
 KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
 KW reperfusion-associated injury; aplastic anaemia; differentiation;
 KW growth; myelodysplastic syndrome; pancytopenic condition;
 KW myocardial ischaemia.

XX Homo sapiens.

XX OS WO9950413-A2.

XX PN 07-OCT-1999.

XX PD 30-MAR-1999; 99WO-US06797.

XX PF 30-MAR-1998; 98US-0079856.

XX PR 20-MAY-1998; 98US-0086074.

Wed Jan 31 13:44:48 2001

```

PR 09-SEP-1998; 98US-0099643.
PR 17-DEC-1998; 98US-0112577.
PR 18-DEC-1998; 98US-0112703.
PR 18-DEC-1998; 98US-0112933.
PR 22-DEC-1998; 98US-0113407.
XX
XX (ELIL ) LILLY & CO ELI.
XX
XX Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
XX PI Hui KY, Khartonenkov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
XX PI Song HY, Wang J, Wu X, Zuckerman SH;
XX
XX WPI; 1999-591319/50.
XX DR N-PSDB; 225377.
XX
XX Use of mature FLINT for treating acute liver failure, inflammation,
XX PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
XX PT and proinflammatory activity
XX
XX Claim 31; Fig 3; 99pp; English.
XX
XX The present invention describes therapeutic applications of mature FLINT
XX (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
XX (mFLINT), which is a member of the tumour necrosis factor receptor
XX superfamily, is used for treating acute liver failure, inflammation of
XX the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
XX with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
XX injury or disorder such as hypercoagulation (including use with
XX thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
XX or disorder, Type I diabetes, cancer, cell damage or damage to an
XX innocent bystander tissue that is induced by a chemotherapeutic agent or
XX therapeutic irradiation, treating haematopoietic progenitor cells that
XX have been exposed to therapeutic radiation or chemotherapy, aplastic
XX anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
XX also used for promoting the growth or differentiation of a haematopoietic
XX progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
XX resulting from abnormal myocardial ischaemia. The present sequence
XX CC represents human mFLINT.
XX
XX Sequence 271 AA;
SQ
Query Match 100.0%; Score 73; DB 20; Length 271;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 THNACRCRTGTF 12
Db 79 thnacrctgtf 90
|||||||

RESULT 6
W66102
ID W66102 standard; Protein; 300 AA.
XX
XX W66102;
XX
XX 02-DEC-1998 (first entry)
XX
XX Amino acid sequence of tumour necrosis related receptor (TR4).
XX Human; tumour necrosis related receptor; TR4; agonist; antagonist;
XX inhibition; chronic; acute; inflammation; arthritis; septicemia;
XX autoimmune disease; transplant rejection; stroke; cancer;
XX Alzheimer's disease.
XX
XX Homo sapiens.
XX
XX EP861850-A1.
XX
XX 02-SEP-1998.
XX
XX 20-JAN-1998; 98EP-0300382.

```

```

XX 04-FEB-1997; 97US-0794796.
XX (SMIK ) SMITHKLINE BEECHAM CORP.
XX
XX Emery J, Tan KB, Truneh A, Young PR;
XX
XX WPI; 1998-508248/44.
XX DR N-PSDB; V07654.
XX
XX New DNA encoding tumour necrosis related receptor - used to treat
XX and prevent e.g. inflammation, arthritis, septicemia, autoimmune
XX diseases, transplant rejection, infection, stroke, ischaemia, ARDS,
XX restenosis, AIDS, bone disorders and cancer
XX
XX Claim 1; Fig 1; 2lpp; English.
XX
XX This is the amino acid sequence of the human tumour necrosis related
XX receptor (TR4), used in the method of the invention. The TR4 protein
XX or its agonist can be used to treat a subject in need of enhanced
XX TR4 polypeptide activity. The antagonist is used to inhibit TR4
XX polypeptide activity. The active agents can be used for the
XX treatment and prevention of diseases such as chronic and acute
XX inflammation, arthritis, septicemia, autoimmune diseases, transplant
XX rejection, stroke, cancer, Alzheimer's disease.
XX
XX Sequence 300 AA;
SQ
Query Match 100.0%; Score 73; DB 19; Length 300;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 THNACRCRTGTF 12
Db 108 thnacrctgtf 119
|||||||

RESULT 7
W63622
ID W63622 standard; Protein; 300 AA.
XX
XX W63622;
XX
XX 26-OCT-1998 (first entry)
XX
XX Human tumour necrosis factor receptor-6 alpha protein.
XX Human tumour necrosis factor receptor-6 alpha; TNFR-6 alpha; TNFR-6 beta;
XX endothelial cells; keratinocytes; normal prostate; apoptosis;
XX prostate tumour tissue.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..30
XX Protein 31..300
XX /note="TNFR-6 alpha"
XX Region 31..282
XX /note="Soluble extracellular domain"
XX
XX W09830694-A2.
XX
XX 16-JUL-1998.
XX
XX 13-JAN-1998; 98WO-US00153.
XX
XX 14-JAN-1997; 97US-0035496.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ebner R, Feng P, Gentz RL, Ni J, Ruben SM, Yu G;
XX
XX

```

DR WPI; 1998-399142/34.
 DR N-PSDB; V39085.
 PT Human tumour necrosis factor receptors 6-alpha and 6-beta - used in
 PT the diagnosis of immune system-related disorder(s)
 XX
 XX Claim 20; Fig 1; 91pp; English.
 PS
 PS The present sequence represents the human tumour necrosis factor
 CC receptor-6 alpha (TNFR-6 alpha) protein. The invention also provides
 CC for the TNFR-6 beta protein (W63623). TNFR-6 alpha and TNFR-6 beta are
 CC members of the tumour necrosis factor receptor (TNFR) family. TNFRs
 CC are expressed in endothelial cells, keratinocytes, normal prostate and
 CC prostate tumour tissue. For a number of disorders of these cells,
 CC particularly of the immune system, substantially altered (whether
 CC increased or decreased) levels of TNFR-6 alpha and/or TNFR-6 beta gene
 CC expression can be detected, therefore the TNFR-6 alpha and TNFR-6 beta
 CC polypeptides, nucleic acids and antibodies are claimed to be useful in
 CC the diagnosis of such disorders. Mutations of the TNFR-6 alpha and
 CC TNFR-6 beta genes can also be detected. The TNFR polypeptides are
 CC also claimed to be useful for identifying ligands which may be useful
 CC in the treatment of apoptosis related disorders.
 XX
 XX SQ Sequence 300 AA;
 Query Match 100.0%; Score 73; DB 19; Length 300;
 Best Local Similarity 100.0%; Pred. No. 0.002;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 THNRACRRTGTF 12
 Db 108 thnracrtgtf 119
 RESULT 8
 Y03099
 ID Y03099 standard; Protein; 300 AA.
 AC Y03099;
 XX
 XX 09-DEC-1999 (first entry)
 DT
 DE Human lung TNF-receptor protein.
 XX
 XX Tumour necrosis factor; TNF; TNF receptor; human; lung; gene therapy;
 KW detection; immunoassay; diagnosis; disease; immune system; tumour;
 KW osteogenic system; cardiovascular system; central nervous system; asthma;
 KW peripheral nervous systems; transplant incompatibility; antitumor;
 KW rheumatoid arthritis; antiasthmatic; antiarthritic.
 XX
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 134..1036
 FT /*tag= a
 FT /*product= "TNF-receptor"
 XX
 PN DE19809978-A1.
 XX
 XX 16-SEP-1999.
 PD
 XX
 XX 09-MAR-1998; 98DE-1009978.
 PF
 XX 09-MAR-1998; 98DE-1009978.
 PR
 XX (BADI) BASF AG.
 PA
 XX Kroeger B;
 PI
 XX WPI; 1999-519473/44.
 DR N-PSDB; Z09998.
 XX

PT New soluble member of tumor necrosis factor receptor family, useful for
 PT identification specific modulators and for treating disease e.g. tumors
 PT
 XX Claim 1; Page 8-9; 10pp; German.
 PS
 PS This invention describes a novel tumour necrosis factor (TNF) receptor
 CC (I) isolated from human lung tissue. (I) is used: (i) to raise specific
 CC antibodies (Ab); (ii) to screen for specific (ant)agonists or ligands
 CC (A), potential therapeutic agents; and (iii) therapeutically (optionally
 CC expressed from a gene therapy vector) in conditions associated with a
 CC deficit of (I). Ab are used: (a) for qualitative or quantitative
 CC detection of (I) in standard immunoassays (for diagnosis of disease, or
 CC susceptibility, or for monitoring); and (b) as therapeutic inhibitors in
 CC cases where (I) is overexpressed. Nucleic acid (II) that encodes (I) is
 CC used: (A) for recombinant production of (I); (B) also its oligonucleotide
 CC fragments, in standard hybridization and/or amplification assays; (C) as
 CC source of antisense molecules or ribozymes; and (D) to produce transgenic
 CC animals (for studying (patho)physiology of (I)). Diseases possibly
 CC associated with under- or over-expression of (I) are those of the immune,
 CC osteogenic, cardiovascular and central or peripheral nervous systems,
 CC tumors, transplant incompatibility, asthma and rheumatoid arthritis. The
 CC products of the invention have antitumor, antiasthmatic and
 CC antiarthritic activity. This sequence represents the TNF-receptor of the
 CC invention.
 XX
 XX SQ Sequence 300 AA;
 Query Match 100.0%; Score 73; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 0.002;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 THNRACRRTGTF 12
 Db 108 thnracrtgtf 119
 RESULT 9
 Y42182
 ID Y42182 standard; Protein; 300 AA.
 AC Y42182;
 XX
 XX 17-DEC-1999 (first entry)
 DT
 DE Human FLINT #1 protein sequence.
 XX
 XX Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
 KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
 KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
 KW reperfusion-associated injury; aplastic anaemia; differentiation;
 KW growth; myelodysplastic syndrome; pancytopenic condition;
 KW myocardial ischaemia.
 XX
 OS Homo sapiens.
 XX
 XX WO9950413-A2.
 PN
 XX 07-OCT-1999.
 PD
 XX
 XX 30-MAR-1999; 99WO-US06797.
 PF
 XX 30-MAR-1998; 98US-0079856.
 PR 20-MAY-1998; 98US-0086074.
 PR 09-SEP-1998; 98US-0099643.
 PR 17-DEC-1998; 98US-0112577.
 PR 18-DEC-1998; 98US-0112703.
 PR 18-DEC-1998; 98US-0112933.
 PR 22-DEC-1998; 98US-0113407.
 XX
 XX (ELIL) LILLY & CO ELI.
 PA
 XX

PI Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
 PI Hui KY, Kharitonov A, Mirzahi J, Na S, Noblitt TW, Reidy CA;
 PI Song HY, Wang J, Wu X, Zuckerman SH;
 XX WPI; 1999-591319/50.
 DR N-PSDB; Z25375.
 XX
 XX Use of mature FLINT for treating acute liver failure, inflammation,
 PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
 PT and proinflammatory activity
 XX
 XX Claim 30; Fig 1; 99pp; English.
 XX
 XX The present invention describes therapeutic applications of mature FLINT
 CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
 CC (mFLINT), which is a member of the tumour necrosis factor receptor
 CC superfamily, is used for treating acute liver failure, inflammation of
 CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
 CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
 CC injury or disorder such as hypercoagulation (including use with
 CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
 CC or disorder, Type I diabetes, cancer, cell damage or damage to an
 CC innocent bystander tissue that is induced by a chemotherapeutic agent or
 CC therapeutic irradiation, treating haematopoietic progenitor cells that
 CC have been exposed to therapeutic radiation or chemotherapy, aplastic
 CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
 CC also used for promoting the growth or differentiation of a haematopoietic
 CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
 CC resulting from abnormal myocardial ischaemia. The present sequence
 CC represents human FLINT.
 XX
 XX Sequence 300 AA;

Query Match 100.0%; Score 73; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 0.002;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 THNRACRRTGF 12
 |||||
 Db 108 thnrcrcrtgf 119

RESULT 10
 Y17479
 ID Y17479 standard; Protein; 300 AA.

XX Y17479;
 AC
 XX
 XX 02-AUG-1999 (first entry)
 XX
 XX Mammalian tumour necrosis factor receptor OPG-2.

XX Tumour necrosis factor receptor; TNF receptor; OPG-2; Paget's disease;
 KW osteopenic disorder; osteoclast activity; primary osteoporosis;
 KW hyperglycaemia; osteolytic metastasis; immune response; cancer.

XX Mammalia.

XX WO9926977-A1.

XX 03-JUN-1999.

XX 24-NOV-1998; 98WO-US25065.

XX 17-FEB-1998; 98US-0074896.

XX 24-NOV-1997; 97US-0066446.

XX (BIOJ) BIOGEN INC.

XX Tschopp J;

XX WPI; 1999-347693/29.

DR N-PSDB; X76052.
 XX
 XX New tumour necrosis factor family receptor OPG-2
 XX
 XX Claim 1; Page 18; 22pp; English.

XX The present sequence represents a mammalian tumour necrosis factor
 CC receptor, designated OPG-2. OPG-2, is a member of the tumour necrosis
 CC factor receptor family, and can be used: (i) to raise specific
 CC antibodies (Ab), (ii) to treat osteopenic disorders associated with
 CC excessive osteoclast activity, e.g. primary osteoporosis, Paget's
 CC disease, hyperglycaemia of malignancy, or osteolytic metastases; (iii)
 CC for affinity purification of cognate ligands, and (iv) to screen for
 CC ligands (antagonists or agonists). Ab, or other OPG-2 blocking agents
 CC such as soluble forms of the protein, are used to prevent, or reduce
 CC severity of, an immune response, and for treating cancer. They can also
 CC be used in diagnostic assays. The nucleic acid sequence encoding OPG-2
 CC can be used as a probe to isolate related sequences from other species.

XX Sequence 300 AA;

Query Match 100.0%; Score 73; DB 20; Length 300;

Best Local Similarity 100.0%; Pred. No. 0.002;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 THNRACRRTGF 12
 |||||
 Db 108 thnrcrcrtgf 119

RESULT 11
 Y06817

ID Y06817 standard; Protein; 300 AA.

XX Y06817;

XX 24-JUN-1999 (first entry)

XX Human Dcr3 polypeptide.

XX Dcr3 polypeptide; tumour necrosis factor receptor; TNFR; Fas ligand;
 KW apoptosis; T cell mediated immune response; allergy; asthma; cancer;
 KW rheumatoid arthritis; Crohn's disease; guest vs. host disease; human;
 KW gene therapy.

XX Homo sapiens.

XX WO9914330-A1.

XX 25-MAR-1999.

XX 18-SEP-1998; 98WO-US19661.

XX 30-JUL-1998; 98US-0094640.

XX 18-SEP-1997; 97US-0059288.

XX (GETH) GENENTECH INC.

XX Ashkenazi AJ, Botstein D, Dodge KH, Goddard A, Gurney AL;

XX Kim KJ, Lawrence DA, Pitti R, Roy MA, Tumas DB;

XX Wood WI;

XX WPI; 1999-244032/20.

XX N-PSDB; X32744.

XX Dcr3 polypeptide related to tumor necrosis factor receptor

XX Claim 5; Fig 1; 88pp; English.

XX This represents a human Dcr3 polypeptide, a homologue of tumour necrosis
 CC factor receptor (TNFR) polypeptide. Host cells containing a vector
 CC comprising the Dcr3 nucleic acid can be used for the recombinant

CC expression of the protein. DcR3 binds to Fas ligand, so it (or its
 CC chimeras) are useful for modulating apoptosis in mammalian cells, also
 CC other Fas-ligand induced activities, particularly to inhibit T cell
 CC mediated immune responses, e.g. in treatment of allergy, asthma,
 CC rheumatoid arthritis, Crohn's disease, guest vs. host disease etc. DcR3
 CC may also be used to identify specific binding proteins, potential
 CC inhibitors. Antibodies against DcR3 are used to treat cancer,
 CC specifically of the lung and colon, also in diagnosis and for affinity
 CC purification of the protein. Detecting mutations in the gene for DcR3 is
 CC also used to diagnose cancer, or predisposition to it. DcR3 nucleic acid
 CC is useful as hybridization probe to detect genomic or related sequences;
 CC for chromosome and gene mapping; as source of antisense sequences; for
 CC expression of recombinant DcR3 and to generate transgenic animals (for
 CC development and screening of therapeutic agents), also for in vivo or
 CC ex vivo gene therapy.

XX Sequence 300 AA;

SQ

Query Match 100.0%; Score 73; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 0.002;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 THNRACRRTGTF 12
 |||||

Db 108 thnrcrctgf 119

RESULT 12

W97749

ID W97749 standard; Protein; 300 AA.

XX AC W97749;

XX DT 21-MAY-1999 (first entry)

XX DE Human tumour necrosis factor receptor ZTNFR-5.

XX KW ZTNFR-5; tumour necrosis factor receptor; TNFR; human;

XX KW cell maturation; bone cell regulation.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Peptide 1..23

FT Protein /note= "signal peptide"

FT Domain 24..300

FT Region /note= "mature protein"

FT Region 24..194

FT Region /note= "extracellular domain"

FT Region 49..71

FT Region /note= "cysteine-rich pseudo-repeat 1"

FT Region 72..113

FT Region /note= "cysteine-rich pseudo-repeat 1"

FT Region 114..151

FT Region /note= "cysteine-rich pseudo-repeat 1"

FT Region 152..194

FT Region /note= "cysteine-rich pseudo-repeat 1"

XX WO9904001-A1.

XX PD 28-JAN-1999.

XX PF 21-JUL-1998; 98WO-US15072.

XX PR 21-JUL-1997; 97US-0053203.

XX PA (ZYMO) ZYMOGENETICS INC.

XX PI Farrah TM;

XX WPI; 1999-132245/11.

XX N-PSDB; X07226.

XX Novel tumour necrosis factor receptor ZTNFR5 - useful for
 PT regulating maturation of TNF-ligand bearing cells

XX Claim 1; Page 84-85; 109pp; English.

XX This polypeptide comprises a new, secreted tumour necrosis factor
 CC receptor (see W97749), designated ZTNFR-5. Novel ZTNFR-5 encoding
 CC polynucleotides and polypeptides were initially identified by
 CC querying an expressed sequence tag (EST) database for sequences
 CC homologous to conserved motifs within the TNF receptor family.
 CC Based on this search, a contig of 16 ESTs (see X07226) was
 CC constructed. ZTNFR-5 polypeptides comprise 4 cysteine-rich repeats
 CC (see also W97750-55) that are homologous to other TNF receptors, in
 CC particular the soluble, secreted TNF receptor osteoprotegerin.
 CC ZTNFR-5 polypeptide can be prepared by recombinant methods. The
 CC polypeptide, especially the extracellular domain, can be used to
 CC generate a soluble variant of ZTNFR-5. The polypeptides and
 CC nucleic acids can be used to screen for ligands, agonists and
 CC antagonists of ZTNFR-5. The polypeptides can be used in bone cell
 CC regulation and to regulate the maturation of TNF ligand-bearing
 CC cells such as T- or B-cells, lymphocytes, peripheral blood
 CC mononuclear cells, polymorphonuclear leukocytes, fibroblasts or
 CC haematopoietic cells.

XX Sequence 300 AA;

SQ

Query Match 100.0%; Score 73; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 0.002;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 THNRACRRTGTF 12
 |||||

Db 108 thnrcrctgf 119

RESULT 13

W95082

ID W95082 standard; Protein; 300 AA.

XX AC W95082;

XX DT 20-MAY-1999 (first entry)

XX DE Orphan receptor (HUMAN NTR-1) polypeptide.

XX KW HUMAN NTR-1; orphan receptor; osteoprotegerin; OPG; TNFR; human;

XX KW tumour necrosis factor receptor; muscle disorder; bone mass; screening;

XX KW muscle metabolism; binding agent; cognate ligand.

XX OS Homo sapiens.

XX PN WO9907738-A2.

XX PD 18-FEB-1999.

XX PF 04-AUG-1998; 98WO-US16202.

XX PR 06-AUG-1997; 97US-0054869.

XX PA (PROC) PROCTER & GAMBLE CO.

XX PA (REG-) REGENERON PHARM INC.

XX PI Masiakowski PJ, Morris J, Valenzuela DM;

XX WPI; 1999-167365/14.

XX N-PSDB; X22300.

XX Novel orphan human receptor polypeptide and nucleic acid - useful as
 PT diagnostic reagents and for treatment of muscle disorders

XX Claim 7; Page 21; 23pp; English.

PS

XX This represents a HUMAN NTR-1 polypeptide, a novel orphan receptor. The
 CC protein is related to osteoprotegerin (OPG) and to tumour necrosis factor
 CC receptor (TNFR). Host cells transformed with a vector comprising the
 CC HUMAN NTR-1 nucleic acid are used for the recombinant expression of the
 CC protein. HUMAN NTR-1 proteins and antibodies immuno specific for the
 CC protein are useful for diagnosis and treatment of humans and animals,
 CC especially muscle disorders, as the receptor is involved in regulation of
 CC bone mass and muscle metabolism. HUMAN NTR-1 receptors are also useful
 CC for screening for novel binding agents, and cognate ligands, which may be
 CC used to treat disorders associated with HUMAN NTR-1 imbalance.

XX Sequence 300 AA;

Query Match 100.0%; Score 73; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 0.002;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 THNRACRCRTGTF 12

Db 108 thnracrcrtg 119

RESULT 14

ID Y77458 standard; Protein; 300 AA.

XX Y77458;

XX 05-JUN-2000 (first entry)

XX Human TNF receptor-like protein, HDTEA84.

XX TNF receptor family; tumour necrosis factor; HDTEA84; HSLJD37R;
 KW Rank-like protein; RANKL; immune disorder; inflammation; allergy;
 KW immunosuppressant; antiarthritic; antirheumatoid; antiinflammatory;
 KW dermatological; antithyroid.

XX Homo sapiens.

XX WO200001817-A2.

XX 13-JAN-2000.

XX 06-JUL-1999; 99WO-US12366.

XX 06-JUL-1998; 98US-0110938.

XX 13-JUL-1998; 98US-0114466.

XX 23-JUL-1998; 98US-0093897.

XX 12-AUG-1998; 98US-0132968.

XX 18-AUG-1998; 98US-0136214.

XX 11-SEP-1998; 98US-0099999.

XX (SCHE) SCHERING CORP.

XX Bates EEM, Lebecque SJE, Murphy EE, Mattson JD, Gorman DM;

PI Hedrick JA, Wang L, Zlotnik A, Murgolo NJ, Greene JR, Johnston JA;

PI Bazan JF, Mahony D, Lees EM;

XX WPI; 2000-171015/15.

DR N-P5DB: 292404.

XX New isolated mammalian genes, used to develop products for treating
 PT e.g. immune, inflammatory or allergic abnormalities, cancers or
 PT degenerative conditions -

XX Claim 24; Page 157; 218pp; English.

XX The invention relates to a number of primate and/or rodent proteins, and
 CC the genes which encode them. The invention encompasses human dendritic
 CC cell prostaglandin transporter (DC-PT); the TNF (tumour necrosis
 CC factor) receptor family-related proteins HDTEA84, HSLJD37R and RANKL;

CC human CC chemokine HCC5; human deubiquitinating proteins Dub1 and Dub
 CC 12; human MD-1 and human and murine MD-2 proteins, which exhibit the
 CC properties of ligands for proteins comprising a leucine-rich motif (LRR);
 CC human cyclin E2; cDNAs encoding these proteins; and antibodies against
 CC these proteins. The proteins can be used for modulating the physiology or
 CC development of a cell. They can be used to mediate uptake of substrates
 CC (e.g., prostaglandin-like molecules), to modulate or mediate cellular
 CC interactions (e.g., induce or prevent trafficking, proliferation, or
 CC differentiation of cells), or are intracellular proteins which are
 CC important in various cellular processes such as the deubiquitination of
 CC proteins or cell cycle regulation. The products can be used for
 CC treating medical conditions such as immune, inflammatory or allergic
 CC disorders, or abnormal cellular proliferation, for example, cancers or
 CC degenerative conditions. They can be used to modulate immune responses in
 CC disease states e.g., autoimmune disorders, including rheumatoid
 CC arthritis, systemic lupus erythematosus, Hashimoto's autoimmune
 CC thyroiditis, as well as acute and chronic inflammatory responses in which
 CC T cell activation, expansion, and/or immunological T cell memory play an
 CC important role. Sequences Y77458-Y77461 and Y77465-Y77468 represent
 CC TNF receptor family-related proteins. Y77458 is the human protein
 CC HDTEA84, Y77459-Y77461 are human HSLJD37R proteins, Y77465 is
 CC murine Rank-like protein RANKL, and Y77466-Y77468 are human RANKL
 CC proteins.

XX Sequence 300 AA;

Query Match 100.0%; Score 73; DB 21; Length 300;

Best Local Similarity 100.0%; Pred. No. 0.002;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 THNRACRCRTGTF 12

Db 108 thnracrcrtg 119

RESULT 15

ID Y42185 standard; Protein; 273 AA.

XX Y42185;

XX 17-DEC-1999 (first entry)

XX Human mFLINT #2 protein sequence.

XX Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
 KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
 KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
 KW reperfusion-associated injury; aplastic anaemia; differentiation;
 KW growth; myelodysplastic syndrome; pancytopenic condition;
 KW myocardial ischaemia.

XX Homo sapiens.

XX WO9950413-A2.

XX 07-OCT-1999.

XX 30-MAR-1999; 99WO-US06797.

XX 30-MAR-1998; 98US-0079856.

XX 20-MAY-1998; 98US-0086074.

XX 09-SEP-1998; 98US-0099643.

XX 17-DEC-1998; 98US-0112577.

XX 18-DEC-1998; 98US-0112703.

XX 18-DEC-1998; 98US-0112933.

XX 22-DEC-1998; 98US-0113407.

XX (ELIL) LILLY & CO ELI.

XX Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;

PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;

PI Song HY, Wang J, Wu X, Zuckerman SH;
XX WPI: 1999-591319/50.
DR N-PSDB; 225378.
XX
PT Use of mature FLINT for treating acute liver failure, inflammation,
PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
PT and proinflammatory activity
XX
PS Example 2; Fig 4; 99pp; English.
XX
CC The present invention describes therapeutic applications of mature FLINT
CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
CC (mFLINT), which is a member of the tumour necrosis factor receptor
CC superfamily, is used for treating acute liver failure, inflammation of
CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
CC injury or disorder such as hypercoagulation (including use with
CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
CC or disorder, Type I diabetes, cancer, cell damage or damage to an
CC innocent bystander tissue that is induced by a chemotherapeutic agent or
CC therapeutic irradiation, treating haematopoietic progenitor cells that
CC have been exposed to therapeutic radiation or chemotherapy, aplastic
CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
CC also used for promoting the growth or differentiation of a haematopoietic
CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
CC resulting from abnormal myocardial ischaemia. The present sequence
CC represents human mFLINT.
XX
SQ Sequence 273 AA;

Query Match 84.9%; Score 62; DB 20; Length 273;
Best Local Similarity 85.7%; Pred. No. 0.064;
Matches 12; Conservative 0; Mismatches 0; Indels 2; Gaps 1;
Qy 1 THNRA--CRCRTGF 12
| | | | | | | | | | | | | |
Db 79 thnrcrcrtgfg 92

Search completed: January 30, 2001, 16:45:30
Job time: 590 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:50:36 ; Search time 149.64 Seconds
(without alignments)
5.445 Million cell updates/sec

Title: US-09-518-931-4_COPY_108_119

Perfect score: 73

Sequence: 1 THNRACRCRTGF 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_66: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	65.8	348	2	D28623
2	48	65.8	349	2	D72175
3	48	65.8	349	2	D36858
4	47	64.4	325	2	B43692
5	47	64.4	3051	2	S42373
6	47	64.4	4543	1	A53102
7	46	63.0	4544	1	S02392
8	46	63.0	4545	1	S25111
9	45	61.6	273	2	PC4133
10	43	58.9	256	2	B32393
11	41	56.2	1193	2	T21133
12	40	54.8	454	1	GOMST1
13	40	54.8	454	2	I57826
14	40	54.8	1429	2	S06434
15	40	54.8	1620	2	T27283
16	39	53.4	277	2	I37552
17	39	53.4	281	2	T24658
18	39	53.4	314	2	I37383
19	39	53.4	324	2	JC2395
20	39	53.4	335	2	A40036
21	39	53.4	461	1	SI8994
22	39	53.4	624	2	T00044
23	39	53.4	788	1	JDVLHH
24	39	53.4	3635	2	T10053
25	39	53.4	5147	1	IJFFPM
26	38.5	52.7	907	1	A57429
27	38	52.1	101	2	F72544
28	38	52.1	210	2	B60030
29	38	52.1	272	2	I48700

ALIGNMENTS

RESULT 1

T28623

hypothetical protein G2R - variola major virus

C:Species: variola major virus

C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 21-Jul-2000

C:Accession: T28623

R:Massung, R.F.; Esposito, J.J.; Liu, L.I.; Qi, J.; Utterback, T.R.; Knight, J.C.; Au

Nature 366, 748-751, 1993

A:Title: Potential virulence determinants in terminal regions of variola smallpox vir

A:Reference number: 220488; MUID:94088747

A:Accession: T28623

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-348 <MAS>

A:Cross-references: EMBL:L22579; NID:G623595; PIDN:AAA60933.1; PID:G439102

A:Experimental source: strain Bangladesh 1975

C:Superfamily: myxoma virus T2 protein; NGF receptor repeat homology

Query Match 65.8%; Score 48; DB 2; Length 348;

Best Local Similarity 58.3%; Pred. No. 1.4;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 THNRACRCRTGF 12

Db 102 THNRICECPGY 113

RESULT 2

D72175

G2R protein - variola minor virus (strain Garcia-1966)

C:Species: variola minor virus

C:Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 20-Jun-2000

C:Accession: D72175

R:Shchelkunov, S.N.; Totmenin, A.V.; Gutorov, V.V.; Safronov, P.F.; Massung, R.F.; Lo

submitted to GenBank, March 1998

A:Description: Analysis of the complete coding sequence of DNA of alastrim variola m

A:Reference number: A72150

A:Accession: D72175

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-349 <SHC>

A:Cross-references: GB:Y16780; NID:G5830555; PIDN:CAB54798.1; PID:G5830759

A:Experimental source: strain Garcia-1966

C:Genetics:

A:Gene: G2R

C:Superfamily: myxoma virus T2 protein; NGF receptor repeat homology

Query Match

Best Local Similarity 58.3%; Pred. No. 1.4;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 THNRACRCRTGF 12
 ||||| | | | |
 Db 103 THNRICECPGY 114

RESULT 3
 D36858
 gene G4R protein - variola virus
 N:Alternate names: B28R protein (COP)
 C:Species: variola virus
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 08-Oct-1999
 C:Accession: D36858; S46888; S32385; S35987
 R:Blinov, V.M.
 submitted to GenBank, November 1992
 A:Description: not shown
 A:Reference number: A36859
 A:Accession: D36858
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-349 <BLI>
 A:Cross-references: GB:X69198; NID:g456758; PIDN:CAA49137.1; PID:g457087
 A:Experimental source: strain India-1967, ssp. major, isolate Ind3
 R:Kolykhalov, A.A.; Blinov, V.M.; Gytarov, V.V.; Pozdnyakov, S.G.; Chizhikov, V.E.; Frolov, V.V.
 submitted to the EMBL Data Library, April 1992
 A:Description: Nucleotide sequence analysis of the region of variola virus XhoI F O H P
 A:Reference number: S46868
 A:Accession: S46888
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-349 <KOL>
 A:Cross-references: EMBL:X67117; NID:g516428; PIDN:CAA47540.1; PID:g516449
 A:Experimental source: strain India-1967, isolate Ind3
 R:Shchelkunov, S.N.; Blinov, V.M.; Sandakchiev, L.S.
 FEBS Lett. 319, 80-83, 1993
 A:Title: Genes of variola and vaccinia viruses necessary to overcome the host protective
 A:Reference number: S32385; MUID:93202281
 A:Accession: S32385
 A:Molecule type: DNA
 A:Residues: 31-168 <SHC>
 A:Cross-references: EMBL:X69198
 A:Experimental source: strain India-1967, ssp. major
 C:Genetics:
 A:Gene: G4R
 C:Superfamily: myxoma virus T2 protein; NGF receptor repeat homology
 F:32-66/Domain: NGF receptor repeat homology <NGF>
 F:68-109/Domain: NGF receptor repeat homology <NG2>
 F:110-151/Domain: NGF receptor repeat homology <NG3>

Query Match 65.8%; Score 48; DB 2; Length 349;
 Best Local Similarity 58.3%; Pred. No. 1.4;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 THNRACRCRTGF 12
 ||||| | | | |
 Db 103 THNRICECPGY 114

RESULT 4
 B43692
 T2 protein - rabbit fibroma virus
 C:Species: rabbit fibroma virus, Shope fibroma virus
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
 C:Accession: B43692
 R:Opton, C.; Delange, A.M.; McFadden, G.
 Virology 160, 20-30, 1987
 A:Title: Tumorigenic poxviruses: genomic organization and DNA sequence of the telomeric
 A:Reference number: A43692; MUID:87321103
 A:Accession: B43692
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-325 <UPT>

A:Cross-references: GB:M17433
 C:Superfamily: myxoma virus T2 protein; NGF receptor repeat homology
 F:64-105/Domain: NGF receptor repeat homology <NG2>
 F:106-147/Domain: NGF receptor repeat homology <NG3>

Query Match 64.4%; Score 47; DB 2; Length 325;
 Best Local Similarity 63.6%; Pred. No. 1.9;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 THNRACRCRTGF 11
 ||||| | | | |
 Db 99 THDRVCNCSTG 109

RESULT 5
 S42373
 hypothetical protein T20G5.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 07-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
 C:Accession: S42373
 R:Smith, A.
 submitted to the EMBL Data Library, March 1994
 A:Reference number: S42368
 A:Accession: S42373
 A:Molecule type: DNA
 A:Residues: 1-3051 <SMI>
 A:Cross-references: EMBL:Z30423; NID:g458479; PID:g458485
 C:Genetics:
 A:introns: 96/1; 166/3; 201/2; 355/1; 404/1; 427/2; 794/1; 892/1; 952/2; 1041/1; 1146
 C:Superfamily: von Willebrand factor type A repeat homology; EGF homology; fibronectin
 F:512-679/Domain: von Willebrand factor type A repeat homology <VWAL>
 F:754-793/Domain: fibronectin type II repeat homology <2F1>
 F:1201-1244/Domain: EGF homology <EGF>

Query Match 64.4%; Score 47; DB 2; Length 3051;
 Best Local Similarity 63.6%; Pred. No. 10;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 HNRACRCRTGF 12
 ||: ||||| | |
 Db 219 HSYSCRCRPGF 229

RESULT 6
 A53102
 alpha-2-macroglobulin receptor precursor - chicken
 N:Alternate names: CD91; LDL receptor-related protein 1; low density lipoprotein rece
 C:Species: Gallus gallus (chicken)
 C:Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 22-Jun-1999
 C:Accession: A53102
 R:Nimpf, J.; Stifani, S.; Bilous, P.T.; Schneider, W.J.
 J. Biol. Chem. 269, 212-219, 1994
 A:Title: The somatic cell-specific low density lipoprotein receptor-related protein o
 A:Reference number: A53102; MUID:94103212
 A:Accession: A53102
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-4543 <NIM>
 A:Cross-references: GB:X74904; NID:g438006; PIDN:CAA52870.1; PID:g438007
 C:Complex: The alpha-2-macroglobulin receptor complex consists of noncovalently-assoc
 d protein.
 C:Superfamily: alpha-2-macroglobulin receptor; EGF homology; LDL receptor ligand-bind
 F:1-17/Domain: signal sequence #status predicted <SIG>
 F:18-3942/Domain: alpha-2-macroglobulin receptor 515K chain #status predicted <S15K>
 F:18-3942,3943-4543/Product: alpha-2-macroglobulin receptor #status predicted <MAT>
 F:29-66/Domain: LDL receptor ligand-binding repeat homology <LDL1>
 F:74-110/Domain: LDL receptor ligand-binding repeat homology <LDL2>
 F:117-150/Domain: EGF homology <EG1>
 F:156-190/Domain: EGF homology <EG2>
 F:200-241/Domain: LDL receptor YWTD-containing repeat homology <YW01>

F: 242~283/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW02>
F: 294~336/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW03>
F: 337~380/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW04>
F: 381~422/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW05>
F: 423~470/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW06>
F: 480~521/Domain:	EGF homology	<EG3>			
F: 573~615/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW07>
F: 616~661/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW08>
F: 662~712/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW09>
F: 713~754/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW10>
F: 755~797/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW11>
F: 805~840/Domain:	EGF homology	<EG4>			
F: 852~888/Domain:	LDL receptor	ligand-binding	repeat	homology	<LDL3>
F: 893~929/Domain:	LDL receptor	ligand-binding	repeat	homology	<LDL4>
F: 934~969/Domain:	LDL receptor	ligand-binding	repeat	homology	<LDL5>
F: 974~1009/Domain:	LDL receptor	ligand-binding	repeat	homology	<LDL6>
F: 1013~1049/Domain:	LDL receptor	ligand-binding	repeat	homology	<LDL7>
F: 1060~1095/Domain:	LDL receptor	ligand-binding	repeat	homology	<LDL8>
F: 1102~1138/Domain:	LDL receptor	ligand-binding	repeat	homology	<LDL9>
F: 1143~1180/Domain:	LDL receptor	ligand-binding	repeat	homology	<LDL10>
F: 1183~1219/Domain:	EGF homology	<EG5>			
F: 1225~1259/Domain:	EGF homology	<EG6>			
F: 1267~1306/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW12>
F: 1307~1353/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW13>
F: 1354~1396/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW14>
F: 1437~1443/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW15>
F: 1444~1486/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW16>
F: 1487~1529/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW17>
F: 1538~1576/Domain:	EGF homology	<EG7>			
F: 1581~1624/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW18>
F: 1626~1667/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW19>
F: 1668~1711/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW20>
F: 1712~1751/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW21>
F: 1752~1794/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW22>
F: 1795~1842/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW23>
F: 1846~1882/Domain:	EGF homology	<EG8>			
F: 1930~1972/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW24>
F: 1973~2015/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW25>
F: 2016~2059/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW26>
F: 2060~2101/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW27>
F: 2102~2147/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW28>
F: 2155~2190/Domain:	EGF homology	<EG9>			
F: 2195~2237/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW29>
F: 2247~2288/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW30>
F: 2338~2382/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW31>
F: 2383~2423/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW32>
F: 2424~2467/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW33>
F: 2476~2511/Domain:	EGF homology	<EG10>			
F: 2518~2555/Domain:	LDL receptor	ligand-binding	repeat	homology	<LDLB>
F: 2560~2594/Domain:	LDL receptor	ligand-binding	repeat	homology	<LDLC>
F: 2599~2633/Domain:	LDL receptor	ligand-binding	repeat	homology	<LDLD>
F: 2646~2682/Domain:	LDL receptor	ligand-binding	repeat	homology	<LDLE>
F: 2690~2724/Domain:	LDL receptor	ligand-binding	repeat	homology	<LDLF>
F: 2732~2767/Domain:	LDL receptor	ligand-binding	repeat	homology	<LDLG>
F: 2772~2810/Domain:	LDL receptor	ligand-binding	repeat	homology	<LDLH>
F: 2816~2851/Domain:	LDL receptor	ligand-binding	repeat	homology	<LDLI>
F: 2862~2895/Domain:	LDL receptor	ligand-binding	repeat	homology	<LDLJ>
F: 2902~2936/Domain:	LDL receptor	ligand-binding	repeat	homology	<LDLK>
F: 2941~2977/Domain:	EGF homology	<EG11>			
F: 2983~3018/Domain:	EGF homology	<EG12>			
F: 3026~3065/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW34>
F: 3066~3110/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW35>
F: 3111~3153/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW36>
F: 3154~3197/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW37>
F: 3198~3238/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW38>
F: 3239~3281/Domain:	LDL receptor	WYTD-containing	repeat	homology	<YW39>
F: 3291~3327/Domain:	EGF homology	<EG13>			
F: 3331~3366/Domain:	LDL receptor	ligand-binding	repeat	homology	<LDL1>
F: 3371~3405/Domain:	LDL receptor	lig			

F:3572-3606/Domain: LDL receptor ligand-binding repeat homology <LDLR>
F:3610-3644/Domain: LDL receptor ligand-binding repeat homology <LDLS>
F:3651-3687/Domain: LDL receptor ligand-binding repeat homology <LDLT>
F:3692-3728/Domain: LDL receptor ligand-binding repeat homology <LDLU>
F:3738-3774/Domain: LDL receptor ligand-binding repeat homology <LDLV>
F:3783-3820/Domain: EGF homology <EG14>
F:3826-3858/Domain: EGF homology <EG15>
F:3866-3908/Domain: LDL receptor WYTD-containing repeat homology <YW40>
F:3910-3968/Domain: LDL receptor WYTD-containing repeat homology <YW41>
F:3943-4543/Domain: alpha-2-macroglobulin receptor 85K chain #status predicted <85K>
F:3943-4420/Domain: 85K chain extracellular #status predicted <EXT>
F:4012-4055/Domain: LDL receptor WYTD-containing repeat homology <YW42>
F:4056-4098/Domain: LDL receptor WYTD-containing repeat homology <YW43>
F:4099-4141/Domain: LDL receptor WYTD-containing repeat homology <YW44>
F:4150-4181/Domain: EGF homology <EG16>
F:4199-4230/Domain: EGF homology <EG17>
F:4235-4266/Domain: EGF homology <EG18>
F:4271-4302/Domain: EGF homology <EG19>
F:4307-4338/Domain: EGF homology <EG20>
F:4343-4373/Domain: EGF homology <EG21>
F:4376-4408/Domain: EGF homology <EG22>
F:4421-4443/Domain: transmembrane #status predicted <TMM>
F:4444-4543/Domain: intracellular #status predicted <INT>
F:116,138,187,276,359,448,731,926,1048,1152,1193,1216,1305,1509,1556,1573,1614,13485,3659,3786,3837,3952,4074,4124,4178,4278/Binding site: carbohydrate (Asn) (covalent)
F:168,2995/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
F:2955/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted

Query 3 NRACRCRTGF 12
DB 503 SRTCRCSGF 512

Query Match 64.4%; Score 47; DB 1; Length 4543;
Best Local Similarity 70.0%; Pred. No. 14;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

RESULT 7
R02392
alpha-2-macroglobulin receptor precursor - human
C:Alternate names: C091; LDL receptor-related protein 1; low density lipoprotein receptor
C:Species: Homo sapiens (man)
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 22-Jun-1999
C:Accession: S02392; S30027; I37998; A39210; S12538
R:Herz, J.; Hamann, U.; Rogne, S.; Myklebost, O.; Gausepohl, H.; Stanley, K.K.
EMBO J. 7, 4119-4127, 1988
A:Title: Surface location and high affinity for calcium of a 500-kd liver membrane protein
A:Reference number: S02392; MUID:89210795
A:Accession: S02392
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-454 <HER>
A:Cross-references: EMBL:X13916; NID:G34338; PIDN:CAA32112.1; PID:G34339
R:Kristensen, T.
submitted to the EMBL Data Library, October 1990
A:Reference number: S30027
A:Accession: S30027
A:Molecule type: mRNA
A:Residues: 3275-3864 <KRI>
A:Cross-references: EMBL:X55077
R:Herz, J.; Kowal, R.C.; Goldstein, J.L.; Brown, M.S.
EMBO J. 9, 1769-1776, 1990
A:Title: Proteolytic processing of the 600 kd low density lipoprotein receptor-related protein
A:Reference number: S12538; MUID:90269210
A:Contents: annotation; site of proteolytic cleavage
R:Kutt, H.; Herz, J.; Stanley, K.K.
Biochim. Biophys. Acta 1009, 229-236, 1989
A:Title: Structure of the low-density lipoprotein receptor-related protein (LRP) from human liver
A:Reference number: I37998; MUID:90089395
A:Accession: I37998
A:Status: preliminary; translated from GB/EMBL/DBJ

A: Molecule type: DNA
A: Residues: 1-11 <RES>
A: Cross-references: EMBL:X15424; NID:q34408; PIDN:CAA33464.1; PTD:g34409
R: Strickland, D.K.; Ashcom, J.D.; Williams, S.; Burgess, W.H.; Migliorini, M.; Argraves, J. Biol. Chem. 265, 17401-17404, 1990
A: Title: Sequence identity between the alpha2-macroglobulin receptor and low density lipoprotein receptor
A: Reference sequence number: A39210; MUID:91009181
A: Accession: A39210
A: Status: preliminary
A: Molecule type: protein
A: Residues: 150-166;234-238, 'X', 240-245, 'X', 247-252; 'G', 686-695;902-916;1096-1109; 'S', 17-18
C: Genetics:
A: Gene: GDB:LRP1; APR; LRP; A2MR
A: Cross-references: GDB:119694; OMIM:107770
A: Map position: 12q13.1-12q13.3
C: Complex: The alpha-2-macroglobulin receptor complex consists of noncovalently-associated protein (see PIR:A39875).
C: Superfamily: alpha-2-macroglobulin receptor; EGF homology; LDL receptor ligand-binding site
C: Keywords: beta-hydroxyasparagine; beta-hydroxyaspartic acid; calcium binding; glycoprotein
F: 1-19/Domain: signal sequence #status predicted <SIG>
F: 20-3943/Product: alpha-2-macroglobulin receptor 515K chain #status predicted <S15K>
F: 27-64/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F: 72-108/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F: 115-148/Domain: EGF homology <EG1>
F: 154-188/Domain: EGF homology <EG2>
F: 198-239/Domain: LDL receptor WYTD-containing repeat homology <YW01>
F: 240-281/Domain: LDL receptor WYTD-containing repeat homology <YW02>
F: 292-334/Domain: LDL receptor WYTD-containing repeat homology <YW03>
F: 335-378/Domain: LDL receptor WYTD-containing repeat homology <YW04>
F: 379-420/Domain: LDL receptor WYTD-containing repeat homology <YW05>
F: 421-468/Domain: LDL receptor WYTD-containing repeat homology <YW06>
F: 478-519/Domain: EGF homology <EG3>
F: 571-613/Domain: LDL receptor WYTD-containing repeat homology <YW07>
F: 614-658/Domain: LDL receptor WYTD-containing repeat homology <YW08>
F: 660-710/Domain: LDL receptor WYTD-containing repeat homology <YW09>
F: 711-752/Domain: LDL receptor WYTD-containing repeat homology <YW10>
F: 753-799/Domain: LDL receptor WYTD-containing repeat homology <YW11>
F: 807-842/Domain: EGF homology <EG4>
F: 854-890/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F: 895-931/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F: 936-971/Domain: LDL receptor ligand-binding repeat homology <LDL5>
F: 976-1011/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F: 1015-1051/Domain: LDL receptor ligand-binding repeat homology <LDL7>
F: 1062-1097/Domain: LDL receptor ligand-binding repeat homology <LDL8>
F: 1104-1140/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F: 1145-1182/Domain: LDL receptor ligand-binding repeat homology <LDL10>
F: 1185-1221/Domain: EGF homology <EG5>
F: 1227-1261/Domain: EGF homology <EG6>
F: 1269-1308/Domain: LDL receptor WYTD-containing repeat homology <YW12>
F: 1309-1355/Domain: LDL receptor WYTD-containing repeat homology <YW13>
F: 1356-1398/Domain: LDL receptor WYTD-containing repeat homology <YW14>
F: 1399-1445/Domain: LDL receptor WYTD-containing repeat homology <YW15>
F: 1446-1488/Domain: LDL receptor WYTD-containing repeat homology <YW16>
F: 1489-1531/Domain: LDL receptor WYTD-containing repeat homology <YW17>
F: 1534-1578/Domain: EGF homology <EG7>
F: 1583-1626/Domain: LDL receptor WYTD-containing repeat homology <YW18>
F: 1627-1669/Domain: LDL receptor WYTD-containing repeat homology <YW19>
F: 1670-1713/Domain: LDL receptor WYTD-containing repeat homology <YW20>
F: 1714-1753/Domain: LDL receptor WYTD-containing repeat homology <YW21>
F: 1754-1796/Domain: LDL receptor WYTD-containing repeat homology <YW22>
F: 1797-1846/Domain: LDL receptor WYTD-containing repeat homology <YW23>
F: 1850-1886/Domain: EGF homology <EG8>
F: 1934-1976/Domain: LDL receptor WYTD-containing repeat homology <YW24>
F: 1977-2019/Domain: LDL receptor WYTD-containing repeat homology <YW25>
F: 2020-2063/Domain: LDL receptor WYTD-containing repeat homology <YW26>
F: 2064-2105/Domain: LDL receptor WYTD-containing repeat homology <YW27>
F: 2106-2151/Domain: LDL receptor WYTD-containing repeat homology <YW28>
F: 2159-2194/Domain: EGF homology <EG9>
F: 2199-2241/Domain: LDL receptor WYTD-containing repeat homology <YW29>
F: 2253-2294/Domain: LDL receptor WYTD-containing repeat homology <YW30>
F: 2344-2388/Domain: LDL receptor WYTD-containing repeat homology <YW31>
F: 2389-2429/Domain: LDL receptor WYTD-containing repeat homology <YW32>
F: 2430-2473/Domain: LDL receptor WYTD-containing repeat homology <YW33>

F: 2482-2517/Domain: EGF homology <EG10>
F: 2524-2561/Domain: LDL receptor ligand-binding repeat homology <LDL10>
F: 2566-2600/Domain: LDL receptor ligand-binding repeat homology <LDL11>
F: 2605-2639/Domain: LDL receptor ligand-binding repeat homology <LDL12>
F: 2652-2688/Domain: LDL receptor ligand-binding repeat homology <LDL13>
F: 2696-2730/Domain: LDL receptor ligand-binding repeat homology <LDL14>
F: 2734-2769/Domain: LDL receptor ligand-binding repeat homology <LDL15>
F: 2774-2812/Domain: LDL receptor ligand-binding repeat homology <LDL16>
F: 2818-2853/Domain: LDL receptor ligand-binding repeat homology <LDL17>
F: 2858-2897/Domain: LDL receptor ligand-binding repeat homology <LDL18>
F: 2904-2939/Domain: LDL receptor ligand-binding repeat homology <LDL19>
F: 2944-2980/Domain: EGF homology <EG11>
F: 2986-3021/Domain: EGF homology <EG12>
F: 3029-3068/Domain: LDL receptor WYTD-containing repeat homology <YW34>
F: 3069-3113/Domain: LDL receptor WYTD-containing repeat homology <YW35>
F: 3114-3156/Domain: LDL receptor WYTD-containing repeat homology <YW36>
F: 3157-3200/Domain: LDL receptor WYTD-containing repeat homology <YW37>
F: 3201-3241/Domain: LDL receptor WYTD-containing repeat homology <YW38>
F: 3242-3284/Domain: LDL receptor WYTD-containing repeat homology <YW39>
F: 3294-3330/Domain: EGF homology <EG13>
F: 3334-3369/Domain: LDL receptor ligand-binding repeat homology <LDL1>
F: 3374-3408/Domain: LDL receptor ligand-binding repeat homology <LDL2>
F: 3413-3448/Domain: LDL receptor ligand-binding repeat homology <LDL3>
F: 3453-3489/Domain: LDL receptor ligand-binding repeat homology <LDL4>
F: 3494-3531/Domain: LDL receptor ligand-binding repeat homology <LDL5>
F: 3536-3570/Domain: LDL receptor ligand-binding repeat homology <LDL6>
F: 3575-3609/Domain: LDL receptor ligand-binding repeat homology <LDL7>
F: 3613-3647/Domain: LDL receptor ligand-binding repeat homology <LDL8>
F: 3654-3690/Domain: LDL receptor ligand-binding repeat homology <LDL9>
F: 3695-3731/Domain: LDL receptor ligand-binding repeat homology <LDL10>
F: 3741-3776/Domain: LDL receptor ligand-binding repeat homology <LDL11>
F: 3785-3822/Domain: EGF homology <EG14>
F: 3828-3860/Domain: EGF homology <EG15>
F: 3868-3911/Domain: LDL receptor WYTD-containing repeat homology <YW40>
F: 3912-3969/Domain: LDL receptor WYTD-containing repeat homology <YW41>
F: 3944-4544/Product: alpha-2-macroglobulin receptor 85K chain #status predicted <85K>
F: 3970-4012/Domain: 85K chain extracellular #status predicted <EXT>
F: 4013-4056/Domain: LDL receptor WYTD-containing repeat homology <YW42>
F: 4057-4099/Domain: LDL receptor WYTD-containing repeat homology <YW43>
F: 4100-4142/Domain: LDL receptor WYTD-containing repeat homology <YW44>
F: 4151-4182/Domain: EGF homology <EG16>
F: 4200-4231/Domain: EGF homology <EG17>
F: 4236-4267/Domain: EGF homology <EG18>
F: 4272-4303/Domain: EGF homology <EG19>
F: 4308-4339/Domain: EGF homology <EG20>
F: 4344-4374/Domain: EGF homology <EG21>
F: 4377-4408/Domain: EGF homology <EG22>
F: 4421-4444/Domain: transmembrane #status predicted <TMM>
F: 4445-4544/Domain: intracellular #status predicted <INT>
F: 166,2998/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
F: 2958/Modified site: erythro-beta-hydroxyaspartic acid (Asp) #status predicted
F: 4075,4125,4278/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 63.0%; Score 46; DB 1; Length 4544;
Best Local Similarity 77.8%; Pred. No. 20;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 RACRCRTGF 12

DB 502 RYCRCSGF 510

RESULT 8

alpha-2-macroglobulin receptor precursor - mouse

N: Alternate names: CD91; LDL receptor-related protein 1; low density lipoprotein rece

C: Species: Mus musculus (house mouse)

C: Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 22-Jun-1999

C: Accession: S25111; S32554

R: van Leuven, F.

submitted to the EMBL Data Library, July 1992

Wed Jan 31 13:44:50 2001

A:Title: Characteristic sequences in the promoter region of the chicken tyrosinase-encoding gene
A:Reference number: PC4153; MUID:96194800
A:Accession: PC4153
A:Molecule type: DNA
A:Residues: 1-273 <FER>
A:Cross-references: GB:L46805; NID:g1146320; PIDN:AAB08441.1; PID:g1146321
A:Note: The DNA sequence includes the first exon and a part of the first intron
C:Comment: This enzyme is the rate-limiting enzyme in pigment biosynthesis.
C:Genetics:
A:Gene: ctyr4.3
C:Superfamily: monophenol monooxygenase
C:Keywords: oxidoreductase

Query Match 61.6%; Score 45; DB 2; Length 273;
Best Local Similarity 63.6%; Pred. No. 3.4;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 HNRACRCRTGF 12
|| |||||
Db 85 YNRCRCRGNF 95

RESULT 10
B32393
T-cell antigen 4-1BB precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 21-Jul-2000
C:Accession: B32393; I48879
R:Kwon, B.S.; Weissman, S.M.
Proc. Natl. Acad. Sci. U.S.A. 86, 1963-1967, 1989
A:Title: cDNA sequence of two inducible T-cell genes.
A:Reference number: A32393; MUID:89184547
A:Accession: B32393
A:Molecule type: mRNA
A:Residues: 1-256 <KWO>
A:Cross-references: GB:J04492; NID:g201121; PIDN:AAA40167.1; PID:g201122
R:Kwon, B.S.; Kozak, C.A.; Kim, K.K.; Pickard, R.T.
J. Immunol. 152, 2256-2262, 1994
A:Title: Genomic Organization and Chromosomal Localization of the T-Cell Antigen 4-1BB.
A:Reference number: I48879; MUID:94179805
A:Accession: I48879
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-256 <RES>
A:Cross-references: EMBL:U02567; NID:g1117783; PIDN:AAA93113.1; PID:g409178
C:Genetics:
A:Introns: 35/1; 115/1; 138/2; 156/3; 181/1; 225/1
C:Superfamily: CD27 antigen; NGF receptor repeat homology
C:Keywords: transmembrane protein
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-256/Product: 4-1BB protein #status predicted <MAT>

Query Match 58.9%; Score 43; DB 2; Length 256;
Best Local Similarity 58.3%; Pred. No. 6.7;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 THNACRCRTGF 12
||| ||||
Db 80 THNAECIEGFG 91

RESULT 11
T21133
hypothetical protein F20B10.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T21133
R:Percy, C.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z19380
A:Accession: T21133

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1193 <WIL>
A:Cross-references: EMBL:Z69636; PIDN:CAA933465.1; GSPDB:GN00022; CESP:F20B10.1
A:Experimental source: clone F20B10
C:Genetics:
A:Gene: CESP:F20B10.1
A:Map position: 4
A:Introns: 18/1; 74/3; 135/1; 222/3; 258/3; 279/1; 496/1; 791/2; 867/1; 969/1;

Query Match 56.2%; Score 41; DB 2; Length 1193;
Best Local Similarity 60.0%; Pred. No. 45;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 HNRACRCRTG 11
||| ||||
Db 654 HNRMCNDSG 663

RESULT 12
COMST1
tumor necrosis factor receptor 1 precursor - mouse
N:Alternate names: tumor necrosis factor receptor, 55K
C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 22-Jun-1999
C:Accession: A38634; B40254; S16677; S19021; I54532
R:Lewis, M.; Tartaglia, L.A.; Lee, A.; Bennett, G.L.; Rice, G.C.; Wong, G.H.W.; Chen, Proc. Natl. Acad. Sci. U.S.A. 88, 2830-2834, 1991
A:Title: Cloning and expression of cDNAs for two distinct murine tumor necrosis factor
A:Reference number: A38634; MUID:91187885
A:Accession: A38634
A:Molecule type: mRNA
A:Residues: 1-454 <LEW>
A:Cross-references: GB:M60468; NID:g199825; PIDN:AAA39751.1; PID:g199826
R:Goodwin, R.G.; Anderson, D.; Jerzy, R.; Davis, T.; Brannan, C.I.; Copeland, N.G.; J Mol. Cell. Biol. 11, 3020-3026, 1991
A:Title: Molecular cloning and expression of the type 1 and type 2 murine receptors f
A:Reference number: A40254; MUID:91246168
A:Accession: B40254
A:Molecule type: mRNA
A:Residues: 1-454 <GO2>
A:Cross-references: GB:M60468; NID:g199825; PIDN:AAA39751.1; PID:g199826
R:Barrett, K.; Taylor-Fishwick, D.A.; Cope, A.P.; Kissonegrhis, A.M.; Gray, P.W.; Fel Eur. J. Immunol. 21, 1649-1656, 1991
A:Title: Cloning, expression and cross-linking analysis of the murine p55 tumor necro
A:Reference number: S16677; MUID:91285014
A:Accession: S16677
A:Molecule type: mRNA
A:Residues: 1-454 <BAR>
A:Cross-references: EMBL:X59238; NID:g53578; PIDN:CAA41922.1; PID:g53579
R:Rothe, J.G.; Brockhaus, M.; Gentz, R.; Lesslauer, W. Immunogenetics 34, 338-340, 1991
A:Title: Molecular cloning and expression of the mouse Tnf receptor type b.
A:Reference number: S19021; MUID:92039815
A:Accession: S19021
A:Molecule type: mRNA
A:Residues: 1-454 <ROT>
A:Cross-references: EMBL:X57796; NID:g54848; PIDN:CAA40936.1; PID:g54849
R:Bebo, B.F. Immunogenetics 39, 450-451, 1994
A:Title: Nucleotide sequence of the TNF type I receptor from a mouse endothelioma cel
A:Reference number: I54532; MUID:94245292
A:Accession: I54532
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-454 <RES>
A:Cross-references: GB:I26349; NID:g430732; PIDN:AAA59361.1; PID:g430733
C:Comment: This protein is one of two distantly related receptors for both TNF-alpha
C:Superfamily: tumor necrosis factor receptor type 1; NGF receptor repeat homology
C:Keywords: duplication; glycoprotein; receptor; transmembrane protein
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-454/Product: tumor necrosis factor receptor type 1 #status predicted <MAT>

Query Match 56.2%; Score 41; DB 2; Length 1193;
Best Local Similarity 60.0%; Pred. No. 45;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 HNRACRCRTG 11
||| ||||
Db 654 HNRMCNDSG 663

F:30-212/Domain: extracellular #status predicted <EXT>
F:44-82/Domain: NGF receptor repeat homology <NG1>
F:84-126/Domain: NGF receptor repeat homology <NG2>
F:127-167/Domain: NGF receptor repeat homology <NG3>
F:168-204/Domain: NGF receptor repeat homology <NG4>
F:213-235/Domain: transmembrane #status predicted <MEM>
F:236-454/Domain: intracellular #status predicted <INT>
F:54,151,202/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 54.8%; Score 40; DB 1; Length 454;
Best Local Similarity 50.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 THNRACRCRTGTF 12
| | | | | | | |
Db 161 TQNTVCNCHAGF 172

RESULT 13

I57826
tumor necrosis factor receptor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 23-Jul-1999
C:Accession: I57826
R:Rothe, J.G.; Bluethmann, H.; Gentz, R.; Lesslauer, W.; Steinmetz, M.
Mol. Immunol. 30, 165-176, 1993
A:Title: Genomic organization and promoter function of the murine tumor necrosis factor
A:Reference number: I57826; MUID:93156721
A:Accession: I57826
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-454 <RES>
A:Cross-references: GB:M76656; NID:g202100; PIDN:AAA0465.1; PID:g202102
C:Genetics:
A:Gene: TNFR-2
A:Introns: 13/3; 65/1; 108/1; 158/1; 184/2; 210/1; 248/1; 257/3; 353/1
C:Superfamily: tumor necrosis factor receptor type 1; NGF receptor repeat homology
C:Keywords: cytokine receptor
F:44-82/Domain: NGF receptor repeat homology <NGF>

Query Match 54.8%; Score 40; DB 2; Length 454;
Best Local Similarity 50.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 THNRACRCRTGTF 12
| | | | | | | |
Db 161 TQNTVCNCHAGF 172

RESULT 14

S06434
homeotic protein lin-12 precursor - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 20-Sep-1999
C:Accession: S06434; A24769
R:Rochem, J.; Weston, K.; Greenwald, I.
Nature 335, 547-550, 1988
A:Title: The Caenorhabditis elegans lin-12 gene encodes a transmembrane protein with ove
A:Reference number: S06434; MUID:88334747
A:Accession: S06434
A:Molecule type: DNA
A:Residues: 1-1429 <YOC>
A:Cross-references: EMBL:MI2069; NID:g156357; PIDN:AAA70191.1; PID:g156358
R:Greenwald, I.
Cell 43, 583-590, 1985
A:Reference number: A24769; MUID:86079540
A:Accession: A24769
A:Molecule type: DNA
A:Residues: 173-712 <GRE>
C:Genetics:
A:Introns: 50/2; 90/1; 109/1; 172/3; 545/1; 589/2; 632/2; 1273/3; 1389/3

C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homol
C:Keywords: glycoprotein; transmembrane protein
F:254-284/Domain: EGF homology <EGF1>
F:507-540/Domain: EGF homology <EGF>
F:547-578/Domain: EGF homology <EGF2>
F:909-931/Domain: transmembrane #status predicted <TMM>
F:1093-1125/Domain: ankyrin repeat homology <ANI>
F:1206-1238/Domain: ankyrin repeat homology <AN2>
F:1240-1272/Domain: ankyrin repeat homology <AN3>

Query Match 54.8%; Score 40; DB 2; Length 1429;
Best Local Similarity 45.5%; Pred. No. 74;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 HNRACRCRTGTF 12
| : | : | : | : | : | :
Db 425 HSPVCQCKNGF 435

RESULT 15

T27283
hypothetical protein Y64G10A.f - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T27283
R:Ainscough, R.
Submitted to the EMBL Data Library, September 1999
A:Reference number: Z20336
A:Accession: T27283
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-1620 <WIL>
A:Cross-references: EMBL:AL110498; NID:el542303; PIDN:CAB54471.1; CESP:Y64G10A.f
A:Gene: CESP:Y64G10A.f
C:Genetics:
A:Experimental source: clone Y64G10A
A:Introns: 77/1; 116/1; 198/1; 282/1; 365/1; 425/1; 466/1; 548/1; 559/1; 601/1; 625/1

Query Match 54.8%; Score 40; DB 2; Length 1620;
Best Local Similarity 85.7%; Pred. No. 82;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 6 CRCRTGTF 12
| | | | | | | |
Db 224 CRCRVGF 230

Search completed: January 30, 2001, 16:50:38
Job time: 728 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: January 30, 2001, 17:06:11 ; Search time 80.04 Seconds
(without alignments)
4.842 Million cell updates/sec

Title: US-09-518-931-4_COPY_108_119
Perfect score: 73
Sequence: 1 THNRACRCRTGF 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	65.8	349	1 VC22_VARV	P34015 variola vir
2	47	64.4	325	1 VT2_SFVKA	P25943 Shope fibro
3	47	64.4	3051	1 YN33_CABEL	P34576 caenorhabdi
4	47	64.4	4543	1 LRPL_CHICK	P98157 gallus gall
5	46	63.0	4544	1 LRPL_HUMAN	Q07954 homo sapien
6	45	61.6	273	1 TYRO_COTJA	Q08410 coturnix co
7	45	61.6	529	1 TYRO_CHICK	P55024 gallus gall
8	43	58.9	256	1 41BB_MOUSE	P20334 mus musculu
9	41	56.2	323	1 FASA_BOVIN	P51867 bos taurus
10	41	56.2	714	1 DLL1_RAT	P97677 rattus norv
11	40	54.8	454	1 TNRL_MOUSE	P25118 mus musculu
12	40	54.8	1429	1 LI12_CABEL	P14585 caenorhabdi
13	39	53.4	277	1 TYRO_TRISI	P55026 trionyx sin
14	39	53.4	273	1 OX40_HUMAN	P43489 homo sapien
15	39	53.4	324	1 FASA_RAT	Q63199 rattus norv
16	39	53.4	332	1 FASA_PIG	O77736 sus scrofa
17	39	53.4	335	1 FASA_HUMAN	P25445 homo sapien
18	39	53.4	461	1 PRTC_RAT	P31394 rattus norv
19	39	53.4	788	1 DPOL_HPBHE	P13846 heron hepat
20	39	53.4	3635	1 LMA5_MOUSE	Q61001 mus musculu
21	39	53.4	5147	1 FAT_DROME	P33450 drosophila
22	38.5	52.7	907	1 MOP_DESGI	Q46509 desulfovibr
23	38	52.1	272	1 OX40_MOUSE	P47741 mus musculu
24	38	52.1	397	1 MPA2_AMBAR	P27762 ambrosia ar
25	38	52.1	461	1 TNRL_HUMAN	P20333 homo sapien
26	38	52.1	515	1 APX1_CABEL	P41990 caenorhabdi
27	38	52.1	672	1 PHX5_MOUSE	P08399 mus musculu
28	38	52.1	740	1 GNT5_RAT	Q08834 rattus norv
29	38	52.1	835	1 CD97_HUMAN	P48960 homo sapien
30	38	52.1	3106	1 LMA2_MOUSE	Q60675 mus musculu
31	38	52.1	3110	1 LMA2_HUMAN	P24043 homo sapien
32	37	50.7	444	1 FA7_RABIT	P98139 oryctolagus
33	37	50.7	466	1 FA7_HUMAN	P08709 homo sapien

34	37	50.7	522	1	TYRL_CARAU	P55028 carassius a
35	37	50.7	532	1	TYRO_RANNI	Q04604 rana nigrom
36	37	50.7	991	1	EMPL_MOUSE	P98063 mus musculu
37	37	50.7	1178	1	TSPL_CHICK	P35440 gallus gall
38	37	50.7	2139	1	CRB_DROME	P10040 drosophila
39	37	50.7	2437	1	NOTC_BRARE	P46530 brachydanio
40	36.5	50.0	712	1	FBL1_CABEL	O77469 caenorhabdi
41	36	49.3	237	1	IPDE_DICDI	P22549 dictyosteli
42	36	49.3	260	1	CD27_HUMAN	P36842 homo sapien
43	36	49.3	396	1	XYRL_YEAST	P46992 saccharomyc
44	36	49.3	413	1	EZF4_HUMAN	Q16254 homo sapien
45	36	49.3	565	1	UBPN_HUMAN	Q9uk80 homo sapien

ALIGNMENTS

```
RESULT 1
VC22_VARV
ID VC22_VARV STANDARD; PRT; 349 AA.
AC P34015;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE PROTEIN C22/B28 HOMOLOG.
GN G4R.
OS Variola virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-INDIA-1967 / ISOLATE IND3;
RX MEDLINE-93202281; PubMed-9384129;
RA Sheelkunov S.N., Blinov V.M., Sandakhchiev L.S.;
RT "Genes of variola and vaccinia viruses necessary to overcome the host
RT protective mechanisms.";
RL FEBS Lett. 319:80-83(1993).
CC -1- SIMILARITY: CONTAINS 2 LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGIONS
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X69198; CAA49137.1; -
CC EMBL; X67117; CAA47540.1; -
CC PIR; D36858; D36858.
CC PIR; S35987; S35987.
CC PIR; S46888; S46888.
CC HSSP; P19438; INCF.
CC INTERPRO; IPR001368; -.
CC PFAM; PF00020; TNFR_C6; 2.
CC PROSITE; PS00652; TNFR_NGFR_1; 2.
CC PROSITE; PS50050; TNFR_NGFR_2; 2.
CC Repeat.
CC DOMAIN 31 108 2 X TNFR-CYS.
CC REPEAT 31 66 TNFR-CYS 1.
CC REPEAT 67 108 TNFR-CYS 2.
CC SEQUENCE 349 AA; 38189 MW; D45D40B5C6E780EF CRC64;
```

Query Match 65.8%; Score 48; DB 1; Length 349;
Best Local Similarity 58.3%; Pred. No. 0.39;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 THNRACRCRTGF 12
Db 103 THNRICEGSPGY 114

```
RESULT 2
VT2_SFVKA STANDARD; PRT; 325 AA.
AC P25943;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE TUMOR NECROSIS FACTOR SOLUBLE RECEPTOR PRECURSOR (PROTEIN T2).
GN T2.
OS Shope fibroma virus (strain Kasza) (SFV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Leporipoxvirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87321103; PubMed=2820128;
RA Upton C., Delange A.M., McFadden G.;
RT "Tumorigenic poxviruses: genomic organization and DNA sequence of the
RL telomeric region of the Shope fibroma virus genome.";
RL Virology 160:20-30(1987).
RN [2]
RP FUNCTION.
RX MEDLINE=91207415; PubMed=1850261;
RA Smith C.A., Davis T., Wignall J.M., Din W.S., Farrah T., Upton C.,
RA McFadden G., Goodwin R.G.;
RT "2 open reading frame from the Shope fibroma virus encodes a soluble
RT form of the TNF receptor.";
RL Biochem. Biophys. Res. Commun. 176:335-342(1991).
CC -1- FUNCTION: BINDS TO TNF-ALPHA AND BETA. PROBABLY PREVENTS TNF TO
CC REACH CELLULAR TARGET AND THEREBY DEAMPENING THE POTENTIAL
CC ANTIVIRAL EFFECTS OF THE CYTOKINE.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M17433; -; NOT_ANNOTATED_CDS.
DR EMBL; A23727; CA01687.1; -;
DR PIR; B43692; B43692.
DR HSP; P19438; 1TNR.
DR INTERPRO: IPR001368; -;
DR PFAM; PF00020; TNFR_C6; 2.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00050; TNFR_NGFR_2; 1.
KW Receptor; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 16
FT CHAIN 17 325 TUMOR NECROSIS FACTOR SOLUBLE RECEPTOR.
FT DOMAIN 27 186 4 X TNFR-CYS.
FT REPEAT 27 62 TNFR-CYS 1.
FT REPEAT 63 104 TNFR-CYS 2.
FT REPEAT 105 147 TNFR-CYS 3.
FT REPEAT 148 186 TNFR-CYS 4.
FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 325 AA; 35132 MW; 810530339198A71E CRC64;

Query Match 64.4%; Score 47; DB 1; Length 325;
Best Local Similarity 63.6%; Pred. NO. 0.53;
Matches 7; Conservative 1; Mismatches 0; Gaps 0;

QY 1 THNRACRRTG 11
Db 99 THDRVCNCSTG 109

RESULT 3
```

```
YNX3_CAEEL STANDARD; PRT; 3051 AA.
ID YNX3_CAEEL
AC P34576;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL PROTEIN T20G5.3 IN CHROMOSOME III (FRAGMENT).
GN T20G5.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Berks M., Smith A.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 45 EGF-LIKE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z30423; CA83007.1; -;
DR PIR; S42373; S42373.
DR HSP; P35555; 1EMO.
DR WORMPEP; T20G5.3; CE00478.
DR INTERPRO: IPR000082; -;
DR INTERPRO: IPR000152; -;
DR INTERPRO: IPR000561; -;
DR INTERPRO: IPR001881; -;
DR INTERPRO: IPR002035; -;
DR PFAM; PF00008; EGF; 31.
DR PFAM; PF01390; SEA; 2.
DR PFAM; PF00092; vwa; 1.
DR PROSITE; PS00010; ASK_HYDROXYL; 28.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 5.
DR PROSITE; PS01187; EGF_CA; 1.
KW Hypothetical protein; EGF-like domain; Repeat; Transmembrane.
FT NON_TER 1 2701
FT DOMAIN <1 2701 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 2702 2722 POTENTIAL.
FT DOMAIN 2723 3051 CYTOPLASMIC (POTENTIAL).
FT DOMAIN <1 44 EGF-LIKE 1 (PARTIAL).
FT DOMAIN 46 94 EGF-LIKE 2.
FT DOMAIN 100 144 EGF-LIKE 3.
FT DOMAIN 145 192 EGF-LIKE 4.
FT DOMAIN 194 245 EGF-LIKE 5.
FT DOMAIN 247 296 EGF-LIKE 6.
FT DOMAIN 313 354 EGF-LIKE 7.
FT DOMAIN 355 402 EGF-LIKE 8.
FT DOMAIN 404 452 EGF-LIKE 9.
FT DOMAIN 454 503 EGF-LIKE 10.
FT DOMAIN 514 599 VWEA.
FT DOMAIN 705 750 EGF-LIKE 11.
FT DOMAIN 750 794 EGF-LIKE 12.
FT DOMAIN 805 846 EGF-LIKE 13.
FT DOMAIN 847 892 EGF-LIKE 14.
FT DOMAIN 892 940 EGF-LIKE 15.
FT DOMAIN 942 990 EGF-LIKE 16.
FT DOMAIN 992 1039 EGF-LIKE 17.
FT DOMAIN 1043 1091 EGF-LIKE 18.
FT DOMAIN 1093 1144 EGF-LIKE 19.
FT DOMAIN 1146 1195 EGF-LIKE 20.
FT DOMAIN 1197 1245 EGF-LIKE 21.
FT DOMAIN 1247 1295 EGF-LIKE 22.
FT DOMAIN 1298 1346 EGF-LIKE 23.
FT DOMAIN 1352 1396 EGF-LIKE 24.
FT DOMAIN 1397 1444 EGF-LIKE 25.
```


DE (ALPHA-2-MACROGLOBULIN RECEPTOR) (A2MR).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WHITE LEGHORN; TISSUE=LIVER, AND OVARY;
RX MEDLINE=94103212; PubMed=7506255;
RA Nimf J., Stifani S., Bilous P.T., Schneider W.J.;
RT "The somatic cell-specific low density lipoprotein receptor-related
RT protein of the chicken. Close kinship to mammalian low density
RT lipoprotein receptor gene family members.";
RL J. Biol. Chem. 269:212-219(1994).
CC -!- FUNCTION: INVOLVED IN THE PLASMA CLEARANCE OF CHYLOMICRON REMNANTS
CC AND ACTIVATED ALPHA 2-MACROGLOBULIN, AS WELL AS THE LOCAL
CC METABOLISM OF COMPLEXES BETWEEN PLASMINOGEN ACTIVATORS AND THEIR
CC ENDOGENOUS INHIBITORS. BINDS VITELLOGENIN, CALCIUM AND ALPHA 2-
CC MACROGLOBULIN.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- ALTERNATIVE PRODUCTS: IN CLONE JN18, AN ASP IS REPLACED BY
CC SER-GLU-ARG-GLN-ASP DUE TO ALTERNATIVE SPLICING OF EXON3.
CC -!- TISSUE SPECIFICITY: SOMATIC.
CC -!- PTM: CLEAVED INTO A 85 KDA MEMBRANE-SPANNING SUBUNIT (LRP-85) AND
CC A 515 KDA LARGE EXTRACELLULAR DOMAIN (LRP-515) THAT REMAINS NON-
CC COVALENTLY ASSOCIATED.
CC -!- SIMILARITY: CONTAINS 22 EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 31 LDL-RECEPTOR CLASS A DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X74904; CAA52870.1; -;
CC HSSP: P01130; 1AJJ.
DR INTERPRO: IPR000033; -;
DR INTERPRO: IPR000152; -;
DR INTERPRO: IPR000561; -;
DR INTERPRO: IPR001438; -;
DR INTERPRO: IPR001881; -;
DR INTERPRO: IPR002172; -;
DR PFAM: PF00008; EGF_16;
DR PFAM: PF00057; ldl_recept_a; 31;
DR PFAM: PF00058; ldl_recept_b; 33;
DR PRINTS: PR00010; EGFBL00D.
DR PRINTS: PR00261; LDLRECEPTOR.
DR PROSITE: PS00010; ASX_HYDROXYL; 3.
DR PROSITE: PS00022; EGF_1; 5.
DR PROSITE: PS01186; EGF_2; 7.
DR PROSITE: PS01187; EGF_CA; 2.
DR PROSITE: PS01209; LDLRA_1; 27.
DR PROSITE: PS03068; LDLRA_2; 31.
DR PROSITE: PS03068; LDLRA_2; 31.
KW Receptor; Transmembrane; Repeat; Endocytosis; Glycoprotein;
KW Signal; Calcium-binding; EGF-like domain; Coated pits;
KW Alternative splicing.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 4543 LOW-DENSITY LIPOPROTEIN RECEPTOR-RELATED
FT PROTEIN 1.
FT DOMAIN 22 4419 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 4420 4443 POTENTIAL.
FT DOMAIN 4444 4543 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 27 68 LDL-RECEPTOR CLASS A 1.
FT DOMAIN 72 112 LDL-RECEPTOR CLASS A 2.
FT DOMAIN 113 151 EGF-LIKE 1.
FT DOMAIN 152 191 EGF-LIKE 2.
FT DOMAIN 476 522 EGF-LIKE 3.
FT DOMAIN 801 841 EGF-LIKE 4.
FT DOMAIN 850 890 LDL-RECEPTOR CLASS A 3.
FT DOMAIN 891 931 LDL-RECEPTOR CLASS A 4.

FT	DOMAIN	932	971	LDL-RECEPTOR CLASS A 5.
FT	DOMAIN	972	1011	LDL-RECEPTOR CLASS A 6.
FT	DOMAIN	1011	1051	LDL-RECEPTOR CLASS A 7.
FT	DOMAIN	1058	1097	LDL-RECEPTOR CLASS A 8.
FT	DOMAIN	1100	1140	LDL-RECEPTOR CLASS A 9.
FT	DOMAIN	1141	1180	LDL-RECEPTOR CLASS A 10.
FT	DOMAIN	1181	1220	EGF-LIKE 5.
FT	DOMAIN	1221	1260	EGF-LIKE 6.
FT	DOMAIN	1534	1577	EGF-LIKE 7.
FT	DOMAIN	1842	1883	EGF-LIKE 8.
FT	DOMAIN	2151	2191	EGF-LIKE 9.
FT	DOMAIN	2472	2512	EGF-LIKE 10.
FT	DOMAIN	2516	2557	LDL-RECEPTOR CLASS A 11.
FT	DOMAIN	2558	2596	LDL-RECEPTOR CLASS A 12.
FT	DOMAIN	2597	2635	LDL-RECEPTOR CLASS A 13.
FT	DOMAIN	2636	2684	LDL-RECEPTOR CLASS A 14.
FT	DOMAIN	2688	2730	LDL-RECEPTOR CLASS A 15.
FT	DOMAIN	2769	2812	LDL-RECEPTOR CLASS A 16.
FT	DOMAIN	2770	2812	LDL-RECEPTOR CLASS A 17.
FT	DOMAIN	2814	2853	LDL-RECEPTOR CLASS A 18.
FT	DOMAIN	2854	2897	LDL-RECEPTOR CLASS A 19.
FT	DOMAIN	2900	2938	LDL-RECEPTOR CLASS A 20.
FT	DOMAIN	2939	2978	EGF-LIKE 11.
FT	DOMAIN	2979	3019	EGF-LIKE 12. CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	3287	3328	EGF-LIKE 13.
FT	DOMAIN	3329	3368	LDL-RECEPTOR CLASS A 21.
FT	DOMAIN	3369	3407	LDL-RECEPTOR CLASS A 22.
FT	DOMAIN	3408	3447	LDL-RECEPTOR CLASS A 23.
FT	DOMAIN	3448	3488	LDL-RECEPTOR CLASS A 24.
FT	DOMAIN	3489	3530	LDL-RECEPTOR CLASS A 25.
FT	DOMAIN	3531	3569	LDL-RECEPTOR CLASS A 26.
FT	DOMAIN	3570	3608	LDL-RECEPTOR CLASS A 27.
FT	DOMAIN	3608	3646	LDL-RECEPTOR CLASS A 28.
FT	DOMAIN	3649	3689	LDL-RECEPTOR CLASS A 29.
FT	DOMAIN	3690	3730	LDL-RECEPTOR CLASS A 30.
FT	DOMAIN	3736	3776	LDL-RECEPTOR CLASS A 31.
FT	DOMAIN	3779	3821	EGF-LIKE 14.
FT	DOMAIN	3822	3859	EGF-LIKE 15.
FT	DOMAIN	4146	4182	EGF-LIKE 16.
FT	DOMAIN	4195	4231	EGF-LIKE 17.
FT	DOMAIN	4231	4267	EGF-LIKE 18.
FT	DOMAIN	4267	4303	EGF-LIKE 19.
FT	DOMAIN	4303	4339	EGF-LIKE 20.
FT	DOMAIN	4339	4374	EGF-LIKE 21.
FT	DOMAIN	4372	4409	EGF-LIKE 22.
FT	SITE	3939	3942	RECOGNITION SITE FOR PROTEOLYTICAL PROCESSING (POTENTIAL).
FT	SITE	4472	4472	CRITICAL FOR ENDOCYTOSIS (BY SIMILARITY).
FT	DOMAIN	4501	4506	ENDOCYTOSIS SIGNAL (POTENTIAL).
FT	DISULFID	29	42	BY SIMILARITY.
FT	DISULFID	36	55	BY SIMILARITY.
FT	DISULFID	49	66	BY SIMILARITY.
FT	DISULFID	74	87	BY SIMILARITY.
FT	DISULFID	81	100	BY SIMILARITY.
FT	DISULFID	94	110	BY SIMILARITY.
FT	DISULFID	117	126	BY SIMILARITY.
FT	DISULFID	122	135	BY SIMILARITY.
FT	DISULFID	137	150	BY SIMILARITY.
FT	DISULFID	156	166	BY SIMILARITY.
FT	DISULFID	162	175	BY SIMILARITY.
FT	DISULFID	177	190	BY SIMILARITY.
FT	DISULFID	480	495	BY SIMILARITY.
FT	DISULFID	491	506	BY SIMILARITY.
FT	DISULFID	508	521	BY SIMILARITY.
FT	DISULFID	805	816	BY SIMILARITY.
FT	DISULFID	812	825	BY SIMILARITY.
FT	DISULFID	827	840	BY SIMILARITY.
FT	DISULFID	852	864	BY SIMILARITY.
FT	DISULFID	859	877	BY SIMILARITY.
FT	DISULFID	871	888	BY SIMILARITY.
FT	DISULFID	893	905	BY SIMILARITY.
FT	DISULFID	900	918	BY SIMILARITY.
FT	DISULFID	912	929	BY SIMILARITY.

FT	DISULFID	934	946	BY SIMILARITY.	RP	SEQUENCE FROM N.A.
FT	DISULFID	941	959	BY SIMILARITY.	RC	TISSUE=LIVER:
FT	DISULFID	953	969	BY SIMILARITY.	RX	MEDLINE=89210795; PubMed=3266596;
FT	DISULFID	974	987	BY SIMILARITY.	RA	Herz J., Hamann U., Rogne S., Myklebost O., Gausepohl H.,
FT	DISULFID	982	1000	BY SIMILARITY.	RA	Stanley K.K.;
FT	DISULFID	994	1009	BY SIMILARITY.	RT	"Surface location and high affinity for calcium of a 500-kd liver
FT	DISULFID	1013	1025	BY SIMILARITY.	RT	membrane protein closely related to the LDL-receptor suggest a
FT	DISULFID	1020	1038	BY SIMILARITY.	RT	physiological role as lipoprotein receptor.";
FT	DISULFID	1032	1049	BY SIMILARITY.	RL	EMBO J. 7:4119-4127(1988).
FT	DISULFID	1060	1073	BY SIMILARITY.	RL	[2]
FT	DISULFID	1067	1086	BY SIMILARITY.	RP	PROCESSING.
FT	DISULFID	1080	1095	BY SIMILARITY.	RX	MEDLINE=90269210; PubMed=2112085;
FT	DISULFID	1102	1116	BY SIMILARITY.	RA	Herz J., Kowal R.C., Goldstein J.L., Brown M.S.;
FT	DISULFID	1120	1129	BY SIMILARITY.	RT	"Proteolytic processing of the 600 kd low density lipoprotein
FT	DISULFID	1123	1138	BY SIMILARITY.	RT	receptor-related protein (LRP) occurs in a trans-Golgi
FT	DISULFID	1143	1157	BY SIMILARITY.	RT	compartment.";
FT	DISULFID	1150	1170	BY SIMILARITY.	RL	EMBO J. 9:1769-1776(1990).
FT	DISULFID	1164	1180	BY SIMILARITY.	RL	[3]
FT	DISULFID	1183	1194	BY SIMILARITY.	RP	FUNCTION.
FT	DISULFID	1190	1204	BY SIMILARITY.	RX	MEDLINE=91092405; PubMed=1702392;
FT	DISULFID	1206	1219	BY SIMILARITY.	RA	Kristensen T., Moestrup S.K., Gliemann J., Bendtsen L., Sand O.,
FT	DISULFID	1225	1235	BY SIMILARITY.	RA	Sotttrup-Jensen L.;
FT	DISULFID	1231	1244	BY SIMILARITY.	RT	"Evidence that the newly cloned low-density-lipoprotein receptor
FT	DISULFID	1246	1259	BY SIMILARITY.	RT	related protein (LRP) is the alpha 2-macroglobulin receptor.";
FT	DISULFID	1538	1551	BY SIMILARITY.	RL	FEBS Lett. 276:151-155(1990).
FT	DISULFID	1547	1561	BY SIMILARITY.	RL	[4]
FT	DISULFID	1563	1576	BY SIMILARITY.	RP	STRUCTURE BY NMR OF 1059-1100.
FT	DISULFID	1846	1857	BY SIMILARITY.	RX	MEDLINE=9253972; PubMed=10318830;
FT	DISULFID	1853	1867	BY SIMILARITY.	RA	Huang W., Dolmer K., Gettins P.G.W.;
FT	DISULFID	1869	1882	BY SIMILARITY.	RT	"NMR solution structure of complement-like repeat CR8 from the low
FT	DISULFID	2155	2166	BY SIMILARITY.	RT	density lipoprotein receptor-related protein.";
FT	DISULFID	2162	2176	BY SIMILARITY.	RL	J. Biol. Chem. 274:14130-14136(1999).
FT	DISULFID	2178	2190	BY SIMILARITY.	CC	-!- FUNCTION: INVOLVED IN THE PLASMA CLEARANCE OF CHYLOMICRON REMNANTS
FT	DISULFID	2476	2487	BY SIMILARITY.	CC	AND ACTIVATED ALPHA 2-MACROGLOBULIN, AS WELL AS THE LOCAL
FT	DISULFID	2483	2497	BY SIMILARITY.	CC	METABOLISM OF COMPLEXES BETWEEN PLASMINOGEN ACTIVATORS AND THEIR
FT	DISULFID	2499	2511	BY SIMILARITY.	CC	ENDOGENOUS INHIBITORS.
FT	DISULFID	2518	2531	BY SIMILARITY.	CC	-!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
FT	DISULFID	2526	2544	BY SIMILARITY.	CC	-!- TISSUE SPECIFICITY: MOST ABUNDANT IN LIVER, BRAIN AND LUNG.
FT	DISULFID	2538	2555	BY SIMILARITY.	CC	-!- PTM: CLEAVED INTO A 85 KDA MEMBRANE-SPANNING SUBUNIT (LRP-85) AND
FT	DISULFID	2560	2572	BY SIMILARITY.	CC	A 515 KDA LARGE EXTRACELLULAR DOMAIN (LRP-515) THAT REMAINS NON-
FT	DISULFID	2567	2585	BY SIMILARITY.	CC	COVALENTLY ASSOCIATED.
FT	DISULFID	2579	2594	BY SIMILARITY.	CC	-!- SIMILARITY: CONTAINS 22 EGF-LIKE DOMAINS.
FT	DISULFID	2599	2611	BY SIMILARITY.	CC	-!- SIMILARITY: CONTAINS 31 LDL-RECEPTOR CLASS A DOMAINS.
FT	DISULFID	2606	2624	BY SIMILARITY.	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
FT	DISULFID	2618	2633	BY SIMILARITY.	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
FT	DISULFID	2638	2660	BY SIMILARITY.	CC	the European Bioinformatics Institute. There are no restrictions on its
FT	DISULFID	2654	2673	BY SIMILARITY.	CC	use by non-profit institutions as long as its content is in no way
FT	DISULFID	2667	2682	BY SIMILARITY.	CC	modified and this statement is not removed. Usage by and for commercial
					CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/
					CC	or send an email to license@isb-sib.ch).
					CC	-----
					DR	EMBL; X13916; CAA32112.1; -
					DR	PDB; 1CR8; 06-JAN-99.
					DR	MIN; 107770; -
					DR	INTERPRO; IPR000033; -
					DR	INTERPRO; IPR000152; -
					DR	INTERPRO; IPR000561; -
					DR	INTERPRO; IPR001881; -
					DR	INTERPRO; IPR002172; -
					DR	PFAM; PF00008; EGF; 16.
					DR	PFAM; PF00057; ldl_recept_a; 31.
					DR	PFAM; PF00058; ldl_recept_b; 33.
					DR	PRINTS; PR00261; LDLRECEPTOR.
					DR	PROSITE; PS00010; ASX_HYDROXYL; 3.
					DR	PROSITE; PS00022; EGF_1; 5.
					DR	PROSITE; PS01186; EGF_2; 8.
					DR	PROSITE; PS01187; EGF_CA; 2.
					DR	PROSITE; PS01209; LDLRA_1; 27.
					DR	PROSITE; PS00068; LDLRA_2; 31.
					KW	Receptor; Transmembrane; Repeat; Endocytosis; Glycoprotein;
					KW	Signal; Calcium-binding; EGF-like domain; Coated pits; 3D-structure.
					FT	SIGNAL 1 19 POTENTIAL.

Query Match 64.4%; Score 47; DB 1; Length 4543;
Best Local Similarity 70.0%; Pred. No. 4.9;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 3 NRACRCRTGF 12
:1111111
Db 503 SRTCCRSGF 512

RESULT 5
LRP1_HUMAN STANDARD; PRT; 4544 AA.
AC Q07954;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE LOW-DENSITY LIPOPROTEIN RECEPTOR-RELATED PROTEIN 1 PRECURSOR (LRP)
DE (ALPHA-2-MACROGLOBULIN RECEPTOR) (A2MR) (APOLOPROTEIN E RECEPTOR)
DE (APOER) (CD91).
GN LRP1 OR A2MR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]

CC THE FORMATION OF PIGMENTS SUCH AS MELANINS AND OTHER POLYPHENOLIC
CC COMPOUNDS. CATALYZES THE RATE-LIMITING CONVERSIONS OF TYROSINE TO
CC DOPA, DOPA TO DOPA-QUINONE AND POSSIBLY 5,6-DIHYDROXYINDOLE TO
CC INDOL-5,6 QUINONE.
CC -1- CATALYTIC ACTIVITY: L-TYROSINE + L-DOPA + O(2) = L-DOPA +
CC DOPAQUINONE + H(2)O.
CC -1- COFACTOR: BINDS TWO COPPER IONS.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MELANOSOMAL.
CC -1- SIMILARITY: BELONGS TO THE TYROSINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S56788; AAB25510.1; -;
DR INTERPRO; IPR002227; -;
DR PFAM; PF00264; Tyrosinase; 1;
DR PROSITE; PS00497; TYROSINASE_1; 1;
DR PROSITE; PS00498; TYROSINASE_2; PARTIAL.
KW Oxidoreductase; Monooxygenase; Copper; Glycoprotein; Signal;
KW Transmembrane; Melanin biosynthesis.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 >273 TYROSINASE.
FT METAL 180 180 COPPER A (BY SIMILARITY).
FT METAL 202 202 COPPER A (BY SIMILARITY).
FT METAL 211 211 COPPER A (BY SIMILARITY).
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 111 111 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 161 161 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NON_TER 273 273
SQ SEQUENCE 273 AA; 31499 MW; 0EA3DE55BE1EA1A CRC64;

Query Match 61.6%; Score 45; DB 1; Length 273;
Best Local Similarity 63.6%; Pred. No. 0.97; Indels 0; Gaps 0;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 HNRACRCRTGF 12
:||| |||||
Db 85 YNRTCRCRGNF 95

RESULT 7
TYRO_CHICK STANDARD; PRT; 529 AA.
ID TYRO_CHICK STANDARD; PRT; 529 AA.
AC P55024;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE TYROSINASE PRECURSOR (EC 1.14.18.1) (MONOPHENOL MONOOXYGENASE).
GN TYR.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WHITE LEGHORN;
RX MEDLINE-93157254; PubMed-1494538;
RA Mochil M., Ito A., Yamamoto H., Takeuchi T., Eguchi G.;
RT "Isolation and characterization of a chicken tyrosinase cDNA.";
RL Pigment Cell Res. 5:162-167(1992).
RN [2]
RP SEQUENCE OF 1-273 FROM N.A.
RX MEDLINE-96194800; PubMed-8647445;
RA Ferguson C.A., Kidson S.H.;
RT "Characteristic sequences in the promoter region of the chicken
RT tyrosinase-encoding gene.";

RL Gene 169:191-195(1996).
CC -1- FUNCTION: THIS IS A COPPER-CONTAINING OXIDASE THAT FUNCTIONS IN
CC THE FORMATION OF PIGMENTS SUCH AS MELANINS AND OTHER POLYPHENOLIC
CC COMPOUNDS. CATALYZES THE RATE-LIMITING CONVERSIONS OF TYROSINE TO
CC DOPA, DOPA TO DOPA-QUINONE AND POSSIBLY 5,6-DIHYDROXYINDOLE TO
CC INDOL-5,6 QUINONE.
CC -1- CATALYTIC ACTIVITY: L-TYROSINE + L-DOPA + O(2) = L-DOPA +
CC DOPAQUINONE + H(2)O.
CC -1- COFACTOR: BINDS TWO COPPER IONS.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MELANOSOMAL.
CC -1- SIMILARITY: BELONGS TO THE TYROSINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D88349; BAA13590.1; -;
DR EMBL; L46805; AAB08441.1; -;
DR INTERPRO; IPR002227; -;
DR PFAM; PF00264; Tyrosinase; 1;
DR PROSITE; PS00092; TYROSINASE.
DR PROSITE; PS00498; TYROSINASE_1; 1;
DR PROSITE; PS00497; TYROSINASE_2; 1;
KW Oxidoreductase; Monooxygenase; Copper; Glycoprotein; Signal;
KW Transmembrane; Melanin biosynthesis.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 529 TYROSINASE.
FT DOMAIN 19 476 LUMENAL, MELANOSOME (POTENTIAL).
FT TRANSMEM 477 497 POTENTIAL.
FT DOMAIN 498 529 CYTOPLASMIC (POTENTIAL).
FT METAL 180 180 COPPER A (BY SIMILARITY).
FT METAL 202 202 COPPER A (BY SIMILARITY).
FT METAL 211 211 COPPER A (BY SIMILARITY).
FT METAL 363 363 COPPER B (BY SIMILARITY).
FT METAL 367 367 COPPER B (BY SIMILARITY).
FT METAL 390 390 COPPER B (BY SIMILARITY).
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 111 111 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 161 161 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 337 337 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 371 371 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 529 AA; 60357 MW; 74B464A52C3EFBF5 CRC64;

Query Match 61.6%; Score 45; DB 1; Length 529;
Best Local Similarity 63.6%; Pred. No. 1.7;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 HNRACRCRTGF 12
:||| |||||
Db 85 YNRTCRCRGNF 95

RESULT 8
41BB_MOUSE STANDARD; PRT; 256 AA.
ID 41BB_MOUSE STANDARD; PRT; 256 AA.
AC P20334;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE 4-LBB LIGAND RECEPTOR PRECURSOR (T-CELL ANTIGEN 4-LBB).
GN TNFRSF9 OR ILA OR LY63 OR CD137 OR CD157.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]

```
RP SEQUENCE FROM N.A.
RX MEDLINE=89184547; PubMed=2784565;
RA Kwon B.S., Weissman S.M.;
RT "cDNA sequences of two inducible T-cell genes.";
RN Proc. Natl. Acad. Sci. U.S.A. 86:1963-1967(1989).
[2]
RP SEQUENCE FROM N.A.
RX STRAIN=BALB/C;
RX MEDLINE=94179805; PubMed=8133039;
RA Kwon B.S., Kozak C.A., Kim K.K., Pickard R.T.;
RT "Genomic organization and chromosomal localization of the T-cell
RT antigen 4-1BB.";
RN J. Immunol. 152:2256-2262(1994).
[3]
RP CHARACTERIZATION, AND SEQUENCE OF 25-29.
RX MEDLINE=93139510; PubMed=7678621;
RA Pollok K.E., Kim Y.-J., Zhou Z., Hurtado J., Kim K.K., Pickard R.T.,
RA Kwon B.S.;
RT "Inducible T cell antigen 4-1BB. Analysis of expression and
RT function.";
RN J. Immunol. 150:771-781(1993).
CC -!- FUNCTION: RECEPTOR FOR THE PUTATIVE CYTOKINE 4-1LBB. POSSIBLY
CC ACTIVE DURING T CELL ACTIVATION.
CC -!- SUBUNIT: PRINCIPALLY AN HOMODIMER, BUT ALSO FOUND AS A MONOMER.
CC ASSOCIATES WITH P56-LCK.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: EXPRESSED ON THE SURFACE OF ACTIVATED T-CELLS.
CC -!- INDUCTION: OPTIMAL BY PMA AND IONOMYCIN.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: J04492; AAA40167.1; -
CC EMBL: U02567; AAA93113.1; -
CC FIR: B32393; B32393.
CC HSSP: P25942; ICDF.
CC
CC INTERPRO: IPR001368; -
CC PFAM: PF00020; TNFR_C6; 1.
CC PROSITE: PS00652; TNFR_NGFR_1; 1.
CC PROSITE: PS50050; TNFR_NGFR_2; FALSE_NEG.
CC Receptor; Glycoprotein; Repeat; Signal; Transmembrane.
KW SIGNAL
FT CHAIN 1 24 4-1BB LIGAND RECEPTOR.
FT DOMAIN 25 256 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 188 208 POTENTIAL.
FT DOMAIN 209 256 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 17 159 4 X TNFR-CYS.
FT REPEAT 17 45 TNFR-CYS 1.
FT REPEAT 46 85 TNFR-CYS 2.
FT REPEAT 86 117 TNFR-CYS 3.
FT REPEAT 118 159 TNFR-CYS 4.
FT CARBOHYD 128 128 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 256 AA; 27598 MW; 93A10D03C60813C4 CRC64;

Query Match 58.9%; Score 43; DB 1; Length 256;
Best Local Similarity 58.3%; Pred. No. 2;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 THNRACRCRTGF 12
Db 80 THNAECCEIEGF 91

RESULT 9
```

```
FASA_BOVIN
ID F51867;
AC 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE FASL RECEPTOR PRECURSOR (APOPTOSIS-MEDIATING SURFACE ANTIGEN FAS)
DE (APO-1 ANTIGEN) (CD95).
GN TNFRSF6 OR APT1 OR FAS.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96226401; PubMed=8634151;
RA Yoo J., Stone R.T., Beattie C.W.;
RT "Cloning and characterization of the bovine Fas.";
RN DNA Cell Biol. 15:227-234(1996).
CC -!- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS FASL. THE
CC ADAPTOR MOLECULE FADD RECRUITS CASPASE-8 TO THE ACTIVATED
CC RECEPTOR. THE RESULTING AGGREGATE CALLED THE DEATH-INDUCING
CC SIGNALING COMPLEX (DISC) PERFORMS CASPASE-8 PROTEOLYTIC
CC ACTIVATION. ACTIVE CASPASE-8 INITIATES THE SUBSEQUENT CASCADE OF
CC CASPASES (ASPARTATE-SPECIFIC CYSTEINE PROTEASES) MEDIATING
CC APOPTOSIS. FAS-MEDIATED APOPTOSIS MAY HAVE A ROLE IN THE
CC INDUCTION OF PERIPHERAL TOLERANCE, IN THE ANTIGEN-STIMULATED
CC SUICIDE OF MATURE T-CELLS, OR BOTH (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- DOMAIN: CONTAINS A DEATH DOMAIN INVOLVED IN THE BINDING OF FADD,
CC AND MAYBE TO OTHER CYTOSOLIC ADAPTOR PROTEINS.
CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -!- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U34794; AAC48546.1; -
CC HSSP: P25445; IDDF.
CC INTERPRO: IPR000488; -
CC INTERPRO: IPR001368; -
CC PFAM: PF00020; TNFR_C6; 3.
CC PFAM: PF00531; Death; 1.
CC PROSITE: PS00652; TNFR_NGFR_1; 2.
CC PROSITE: PS50050; TNFR_NGFR_2; 2.
CC PROSITE: PS50017; DEATH_DOMAIN; 1.
CC Apoptosis; Receptor; Glycoprotein; Transmembrane; Repeat; Signal.
KW SIGNAL
FT CHAIN 1 16 POTENTIAL.
FT DOMAIN 17 323 FASL RECEPTOR.
FT DOMAIN 17 170 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 171 188 POTENTIAL.
FT DOMAIN 189 323 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 45 163 3 X TNFR-CYS.
FT REPEAT 45 80 TNFR-CYS 1.
FT REPEAT 81 124 TNFR-CYS 2.
FT REPEAT 125 163 TNFR-CYS 3.
FT REPEAT 238 306 DEATH DOMAIN.
SQ SEQUENCE 323 AA; 36445 MW; 4D88A90E9E1F4892 CRC64;
```

```
Query Match 56.2%; Score 41; DB 1; Length 323;
Best Local Similarity 50.0%; Pred. No. 5.1;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 THNRACRCRTGF 12
Db 119 TRNTKCRCKSNF 130
```

```

RESULT 10
DLL1_RAT          STANDARD;          PRT;    714 AA.
ID  P97677;
AC  01-NOV-1997 (Rel. 35, Created)
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DT  15-JUL-1999 (Rel. 38, Last annotation update)
DE  DELTA-LIKE PROTEIN 1 PRECURSOR (DELTA1).
GN  DLL1.
OS  Rattus norvegicus (Rat).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Disibio G., Hebshi L., Boulter J., Weinmaster G.;
RL  Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
CC  -!- FUNCTION: MAY BE INVOLVED IN CELL-TO-CELL COMMUNICATION IN
CC  MAMMALIAN EMBRYOS. MAY HAVE A ROLE IN CELLULAR INTERACTIONS
CC  UNDERLYING SOMITOGENESIS AND DEVELOPMENT OF THE NERVOUS SYSTEM (BY
CC  SIMILARITY).
CC  -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC  -!- SIMILARITY: CONTAINS 8 EGF-LIKE DOMAINS.
CC  -!- SIMILARITY: BELONGS TO THE DELTA/SERRATE/JAGGED FAMILY.
-----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
-----
EMBL; U78889; AA837343.1; -
DR  HSP; P00740; IIXA.
DR  INTERPRO; IPR000152; -
DR  INTERPRO; IPR000561; -
DR  INTERPRO; IPR001438; -
DR  INTERPRO; IPR001774; -
DR  INTERPRO; IPR001881; -
DR  PFAM; PF01414; DSL; 1.
DR  PFAM; PF00008; EGF; 7.
DR  PRINTS; PR00010; EGFBL0D.
DR  PROSITE; PS00010; ASX HYDROXYL; 3.
DR  PROSITE; PS00022; EGF_1; 8.
DR  PROSITE; PS01186; EGF_2; 8.
DR  PROSITE; PS01187; EGF_CA; 2.
KW  Signal; EGF-like domain; Glycoprotein; Transmembrane.
FT  SIGNAL 1
FT  CHAIN 17
FT  CHAIN 18 714
FT  DOMAIN 18 537
FT  TRANSMEM 538
FT  DOMAIN 561 714
FT  DOMAIN 225 253
FT  DOMAIN 256 284
FT  DOMAIN 291 324
FT  DOMAIN 331 362
FT  DOMAIN 369 401
FT  DOMAIN 408 439
FT  DOMAIN 446 477
FT  DOMAIN 484 515
FT  DISULFID 225 236
FT  DISULFID 229 242
FT  DISULFID 244 253
FT  DISULFID 256 267
FT  DISULFID 262 273
FT  DISULFID 275 284
FT  DISULFID 291 303
FT  DISULFID 297 313
FT  DISULFID 315 324
FT  DISULFID 331 342
FT  DISULFID 336 351
FT  DISULFID 353 362

```

```

FT  DISULFID 369 380
FT  DISULFID 374 390
FT  DISULFID 392 401
FT  DISULFID 408 419
FT  DISULFID 413 428
FT  DISULFID 430 439
FT  DISULFID 446 466
FT  DISULFID 468 477
FT  DISULFID 484 495
FT  DISULFID 489 504
FT  DISULFID 506 515
FT  CARBOHYD 476 476
SQ  SEQUENCE 714 AA; 77378 MW; 4B8EE2272BAEA27E CRC64;

Query Match          56.2%; Score 41; DB 1; Length 714;
Best Local Similarity 85.7%; Pred. No. 9.9;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY  6 CRCRTGF 12
    III:III
DB  428 CRCQTGF 434

RESULT 11
TNRL_MOUSE
ID  TNRL_MOUSE          STANDARD;          PRT;    454 AA.
AC  P25118;
DT  01-MAY-1992 (Rel. 22, Created)
DT  01-MAY-1992 (Rel. 22, Last sequence update)
DT  01-OCT-2000 (Rel. 40, Last annotation update)
DE  TUMOR NECROSIS FACTOR RECEPTOR 1 PRECURSOR (P60) (TNF-R1) (TNF-RI)
DE  (P55)
GN  TNFRSF1A OR TNFR1 OR TNFR-1.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE-91187885; PubMed-1849278;
RA  Lewis M., Tartaglia L.A., Lee A., Bennett G.L., Rice G.C.,
RA  Wong G.H., Chen E.Y., Goeddel D.V.;
RT  "Cloning and expression of cDNAs for two distinct murine tumor
RT  necrosis factor receptors demonstrate one receptor is species
RT  specific."
RT  Proc. Natl. Acad. Sci. U.S.A. 88:2830-2834(1991).
RN  [2]
RP  SEQUENCE FROM N.A.
RX  MEDLINE-91246168; PubMed-1645445;
RA  Goodwin R.G., Anderson D., Jerzy R., Davis T., Brannan C.I.,
RA  Copeland N.G., Jenkins N.A., Smith C.A.;
RT  "Molecular cloning and expression of the type 1 and type 2 murine
RT  receptors for tumor necrosis factor."
RL  Mol. Cell. Biol. 11:3020-3026(1991).
RN  [3]
RP  SEQUENCE FROM N.A.
RX  MEDLINE-91285014; PubMed-1647956;
RA  Barrett K., Taylor-Fishwick D.A., Cope A.P., Kissenerghis A.M.,
RA  Gray P.W., Feldmann M., Foxwell B.M.J.;
RT  "Cloning, expression and cross-linking analysis of the murine p55
RT  tumor necrosis factor receptor."
RL  Eur. J. Immunol. 21:1649-1656(1991).
RN  [4]
RP  SEQUENCE FROM N.A.
RX  MEDLINE-92039815; PubMed-1657766;
RA  Rothe J.G., Brockhaus M., Gentz R., Lesslauer W.;
RT  "Molecular cloning and expression of the mouse Tnf receptor type b."
RL  Immunogenetics 34:338-340(1991).
RN  [5]
RP  SEQUENCE FROM N.A.
RX  MEDLINE-94245292; PubMed-8188324;
RA  Bebo B.F., Linthicum D.S.;

```

RT "Nucleotide sequence of the TNF type I receptor from a mouse
 RT endothelioma cell line."
 RL Immunogenetics 39:450-451(1994).
 RN [6]
 RX SEQUENCE FROM N.A.
 RA MEDLINE=93156721; PubMed=8381516;
 RA Rothe J., Bluethmann H., Gentz R., Lesslauer W., Steinmetz M.;
 RT "Genomic organization and promoter function of the murine tumor
 RT necrosis factor receptor beta gene."
 RL Mol. Immunol. 30:165-175(1993).
 CC -!- FUNCTION: RECEPTOR FOR TNF-ALPHA. THE ADAPTOR MOLECULE FADD
 CC RECRUITS CASPASE-8 TO THE ACTIVATED RECEPTOR. THE RESULTING
 CC AGGREGATE CALLED THE DEATH-INDUCING SIGNALING COMPLEX (DISC)
 CC PERFORMS CASPASE-8 PROTEOLYTIC ACTIVATION WHICH INITIATES THE
 CC SUBSEQUENT CASCADE OF CASPASES (ASPARTATE-SPECIFIC CYSTEINE
 CC PROTEASES) MEDIATING APOPTOSIS (BY SIMILARITY).
 CC -!- SUBUNIT: TNF BINDING TO THE EXTRACELLULAR DOMAIN OF TNFR1 LEADS TO
 CC HOMOTRIMERIZATION. ONCE AGGREGATED THE RECEPTORS DEATH DOMAINS
 CC PROVIDE A NOVEL MOLECULAR INTERFACE THAT INTERACTS SPECIFICALLY
 CC WITH THE DEATH DOMAIN OF TRADD. VARIOUS TRADD-INTERACTING
 CC PROTEINS SUCH AS TRAFs, RIP AND POSSIBLY FADD, ARE RECRUITED TO
 CC TNFR1 COMPLEX BY THEIR ASSOCIATION WITH TRADD. THIS COMPLEX
 CC ACTIVATES AT LEAST TWO DISTINCT SIGNALING CASCADES, APOPTOSIS AND
 CC NF-KAPPA B SIGNALING (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 CC -!- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; M60468; AAA39751.1; -
 DR EMBL; M59377; AAA40464.1; -
 DR EMBL; X59238; AAA41922.1; -
 DR EMBL; X57796; CAA40936.1; -
 DR EMBL; L26349; CAA59361.1; -
 DR EMBL; M76656; AAA40465.1; -
 DR EMBL; M88067; AAA40465.1; JOINED.
 DR EMBL; M76635; AAA40465.1; JOINED.
 DR PIR; A38634; GQMST1.
 DR PIR; S16677; S16677.
 DR PIR; S19021; S19021.
 DR HSSP; P19438; LEXT.
 DR MGD; MGI:1314884; TNFRSF1A.
 DR INTERPRO; IPR000488; -
 DR INTERPRO; IPR001368; -
 DR PFAM; PF00020; TNFR_C6; 4.
 DR PFAM; PF00531; death; 1.
 DR PROSITE; PS00652; TNFR_NGFR_1; 3.
 DR PROSITE; PS50050; TNFR_NGFR_2; 3.
 DR PROSITE; PS50017; DEATH_DOMAIN; 1.
 KW Receptor; Transmembrane; Glycoprotein; Repeat; Signal; Apoptosis.
 FT SIGNAL 1 21
 FT CHAIN 22 454
 FT DOMAIN 22 212
 FT TRANSMEM 213 235
 FT DOMAIN 236 454
 FT DOMAIN 43 196
 FT REPEAT 43 82
 FT REPEAT 83 125
 FT REPEAT 126 166
 FT REPEAT 167 196
 FT DOMAIN 339 349
 FT DOMAIN 356 441
 FT DISULFID 44 58
 FT DISULFID 59 72
 FT DISULFID 62 81
 FT DISULFID 84 99
 BY SIMILARITY.

FT DISULFID 102 117 BY SIMILARITY.
 FT DISULFID 125 125 BY SIMILARITY.
 FT DISULFID 127 143 BY SIMILARITY.
 FT DISULFID 146 158 BY SIMILARITY.
 FT DISULFID 149 166 BY SIMILARITY.
 FT DISULFID 168 179 BY SIMILARITY.
 FT DISULFID 182 191 BY SIMILARITY.
 FT DISULFID 185 195 BY SIMILARITY.
 FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 202 202 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 394 394 R -> G (IN REF. 6).
 SQ SEQUENCE 454 AA; 50129 MW; 0710C2E8C3C2B6D9 CRC64;
 Query Match 54.8%; Score 40; DB 1; Length 454;
 Best Local Similarity 50.08; Pred. No. 9.8;
 Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 THNRACRCRTGF 12
 Db 161 TQNTVCNCHAGF 172
 ID LI12 CAEEL STANDARD; PRT; 1429 AA.
 AC P14585;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE LIN-12 PROTEIN PRECURSOR.
 GN LIN-12 OR R107.8.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=88334747; PubMed=3419531;
 RA Yochem J., Weston K., Greenwald I.;
 RT "The Caenorhabditis elegans lin-12 gene encodes a transmembrane
 RL protein with overall similarity to Drosophila Notch."
 RN Nature 335:547-550(1988).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fraser A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
 RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
 RA Sims M., Smalton N., Smith A., Smith M., Sonhammer E., Staden K.,
 RA Sulston J., Thierly-Mieg J., Thomas K., Vaudin M., Vaughan K.,
 RA Watson R., Watson A., Weinstock L., Wilkinson-Sproat J.,
 RA Wooldman P.;
 RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans."
 RL Nature 368:32-38(1994).
 CC -!- FUNCTION: LIN-12 IS INVOLVED IN SEVERAL CELL FATES DECISIONS THAT
 CC REQUIRES CELL-CELL INTERACTIONS. IT IS POSSIBLE THAT LIN-12
 CC ENCODES A MEMBRANE-BOUND RECEPTOR FOR A SIGNAL THAT ENABLES
 CC EXPRESSION OF THE VENTRAL UTERINE PRECURSOR CELL FATE.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -!- SIMILARITY: HIGH, TO C.ELEGANS GLP-1.
 CC -!- SIMILARITY: CONTAINS 13 EGF-LIKE DOMAINS.
 CC -!- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
 CC -!- SIMILARITY: CONTAINS 6 ANK REPEATS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; M12069; AAA70191.1; -;
 DR EMBL; Z14092; CAA78474.1; -;
 DR PIR; S06434; S06434.
 DR HSP; P00740; LIXA.
 DR WORMPEP; R107.8; CE00274.
 DR INTERPRO; IPR000152; -;
 DR INTERPRO; IPR000561; -;
 DR INTERPRO; IPR000800; -;
 DR INTERPRO; IPR001881; -;
 DR INTERPRO; IPR002110; -;
 DR PFAM; PF00008; EGF_13;
 DR PFAM; PF00023; ank; 4;
 DR PFAM; PF00066; notch; 3;
 DR PROSITE; PS50088; ANK_REPEAT; 3;
 DR PROSITE; PS50297; ANK_REPEAT_REGION; 1;
 DR PROSITE; PS00010; ASX_HYDROXYL; 3;
 DR PROSITE; PS00022; EGF_1; 12;
 DR PROSITE; PS01186; EGF_2; 11;
 DR PROSITE; PS01187; EGF_CA; 2;
 DR Differentiation; Repeat; ANK repeat; EGF-like domain; Transmembrane;
 KW Glycoprotein; Signal;
 FT SIGNAL 1
 FT CHAIN 16 1429
 FT DOMAIN 16 908
 FT TRANSMEM 909 931
 FT DOMAIN 932 1429
 FT DOMAIN 24 618
 FT DOMAIN 631 750
 FT DOMAIN 1046 1266
 FT DOMAIN 20 61
 FT DOMAIN 114 150
 FT DOMAIN 152 190
 FT DOMAIN 201 246
 FT DOMAIN 250 285
 FT DOMAIN 287 323
 FT DOMAIN 323 363
 FT DOMAIN 365 402
 FT DOMAIN 404 441
 FT DOMAIN 449 492
 FT DOMAIN 503 541
 FT DOMAIN 543 579
 FT DOMAIN 582 619
 FT REPEAT 635 669
 FT REPEAT 670 710
 FT REPEAT 711 750
 FT REPEAT 1046 1078
 FT REPEAT 1079 1119
 FT REPEAT 1120 1152
 FT REPEAT 1153 1188
 FT REPEAT 1189 1232
 FT REPEAT 1233 1266
 FT DISULFID 24 35
 FT DISULFID 29 49
 FT DISULFID 51 60
 FT DISULFID 118 129
 FT DISULFID 123 138
 FT DISULFID 140 149
 FT DISULFID 156 169
 FT DISULFID 163 178
 FT DISULFID 180 189
 FT DISULFID 205 227
 FT DISULFID 221 234
 FT DISULFID 236 245
 FT DISULFID 254 264
 FT DISULFID 259 273
 FT DISULFID 275 284

FT DISULFID 291 302 BY SIMILARITY.
 FT DISULFID 296 311 BY SIMILARITY.
 FT DISULFID 313 322 BY SIMILARITY.
 FT DISULFID 327 339 BY SIMILARITY.
 FT DISULFID 334 351 BY SIMILARITY.
 FT DISULFID 353 362 BY SIMILARITY.
 FT DISULFID 369 381 BY SIMILARITY.
 FT DISULFID 375 390 BY SIMILARITY.
 FT DISULFID 392 401 BY SIMILARITY.
 FT DISULFID 408 419 BY SIMILARITY.
 FT DISULFID 413 429 BY SIMILARITY.
 FT DISULFID 431 440 BY SIMILARITY.
 FT DISULFID 507 518 BY SIMILARITY.
 FT DISULFID 512 529 BY SIMILARITY.
 FT DISULFID 531 540 BY SIMILARITY.
 FT DISULFID 547 558 BY SIMILARITY.
 FT DISULFID 552 567 BY SIMILARITY.
 FT DISULFID 569 578 BY SIMILARITY.
 FT DISULFID 586 597 BY SIMILARITY.
 FT DISULFID 591 607 BY SIMILARITY.
 FT DISULFID 609 618 BY SIMILARITY.
 FT CARBOHYD 41 41 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 165 165 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 194 194 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 378 378 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 515 515 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 623 623 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 751 751 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 754 754 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 900 900 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 1429 AA; 157115 MW; 255EDD7A62C025DB CRC64;
 Query Match 54.8%; Score 40; DB 1; Length 1429;
 Best Local Similarity 45.5%; Pred. No. 26;
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 2 HNRACRCRTGF 12
 I: |::||
 Db 425 HSPVCQCKNGF 435
 RESULT 13
 TYRO_TRISI STANDARD; PRT; 273 AA.
 AC P55026;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE TYROSINASE PRECURSOR (EC 1.14.18.1) (MONOPHENOL MONOOXYGENASE)
 DE (FRAGMENT).
 GN TYR.
 OS Trionyx sinensis (Chinese softshell turtle) (Pelodiscus sinensis).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Testudines; Cryptodira; Trionychoidae; Trionychidae; Trionyx.
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SSP. JAPONICUS; TISSUE-TESTIS;
 RX MEDLINE=93181407; Pubmed=1292011;
 RA Yamamoto H., Kudo T., Masuko N., Miura H., Sato S., Tanaka M.,
 RA Tanaka S., Takeuchi S., Shibahara S., Takeuchi T.;
 RT "Phylogeny of regulatory regions of vertebrate tyrosinase genes.";
 RL Pigment Cell Res. 5:284-294(1992).
 CC -!- FUNCTION: THIS IS A COPPER-CONTAINING OXIDASE THAT FUNCTIONS IN
 CC THE FORMATION OF PIGMENTS SUCH AS MELANINS AND OTHER POLYPHENOLIC
 CC COMPOUNDS.
 CC -!- CATALYTIC ACTIVITY: L-TYROSINE + L-DOPA + O(2) = L-DOPA +
 CC DOPAQUINONE + H(2)O.
 CC -!- COFACTOR: BINDS TWO COPPER IONS.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. MELANOSOMAL.
 CC -!- SIMILARITY: BELONGS TO THE TYROSINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; S56789; AAB25511.1; -
HSP; P02468; L1LE.
INTERPRO: IPR002227; -
PFAM; PF00264; TYROSINASE; 1.
PROSITE; PS00497; TYROSINASE_1;
PROSITE; PS00498; TYROSINASE_2; PARTIAL.
Oxidoreductase; Monooxygenase; Copper; Glycoprotein; Signal;
Transmembrane; Melanin biosynthesis.
KW SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 >273 TYROSINASE.
FT METAL 180 COPPER A (BY SIMILARITY).
FT METAL 202 202 COPPER A (BY SIMILARITY).
FT METAL 211 211 COPPER A (BY SIMILARITY).
FT CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 169 169 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NON_TER 273 273
SEQUENCE 273 AA; 31136 MW; 9365C76EDB26A625 CRC64;

Query Match 53.4%; Score 39; DB 1; Length 273;
Best Local Similarity 45.5%; Pred. No. 9.3;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 HNRACRGTGF 12
DB 85 INRTCQCSNF 95

RESULT 14
OX40_HUMAN STANDARD; PRT; 277 AA.
ID OX40_HUMAN
AC P43489; Q13663;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE OX40L RECEPTOR PRECURSOR (ACT35 ANTIGEN) (TAX-TRANSCRIPTIONALLY
DE ACTIVATED GLYCOPROTEIN 1 RECEPTOR) (CD134 ANTIGEN).
GN TNFRSF4 OR TXGPIL.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
[1]
SEQUENCE FROM N.A.
MEDLINE-94170844; PubMed-7510240;
Latz U., Duerkop H., Schnittger S., Ringeling J., Eitelbach F.,
Hummel M., Fonatsch C., Stein H.;
"The human OX40 homolog: cDNA structure, expression and chromosomal
assignment of the ACT35 antigen."
Eur. J. Immunol. 24:677-683(1994).
[2]
SEQUENCE FROM N.A.
MEDLINE-95219871; PubMed-7704935;
Baum P.R., Gayle R.B. III, Ramsdell F., Srinivasan S., Sorensen R.A.,
Watson M.L., Seidlin M.F., Clifford K.N., Grabstein K., Alderson M.R.;
"Identification of OX40 ligand and preliminary characterization of
its activities on OX40 receptor."
Circ. Shock 44:30-34(1994).
RL Circ Shock 44:30-34(1994).
CC -!- FUNCTION: RECEPTOR FOR THE OX40L/GP34 CYTOKINE.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- SIMILARITY: CONTAINS A LA-NGR/TFNFR-TYPE CYSTEINE-RICH REGION.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide cd134 entry;
WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd134.htm".
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -

EMBL; X75962; CAA33576.1; -
EMBL; S76792; AAB33944.1; ALT_INIT.
HSP; P25942; 1CDF.
MIM; 600315; -
INTERPRO: IPR001368; -
PFAM; PF00020; TNFR_C6; 3.
PROSITE; PS00652; TNFR_NGFR_1; 3.
PROSITE; PS00050; TNFR_NGFR_2; 2.
Receptor; T-cell; Antigen; Glycoprotein; Transmembrane; Repeat;
KW SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 277 OX40L RECEPTOR.
FT DOMAIN 29 214 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 215 235 POTENTIAL.
FT DOMAIN 236 277 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 30 169 4 X TNFR-CYS.
FT REPEAT 30 65 TNFR-CYS 1.
FT REPEAT 66 107 TNFR-CYS 2.
FT REPEAT 108 126 TNFR-CYS 3 (INCOMPLETE).
FT REPEAT 127 167 TNFR-CYS 4.
FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 277 AA; 29340 MW; 49F15525941550BF CRC64;

Query Match 53.4%; Score 39; DB 1; Length 277;
Best Local Similarity 54.5%; Pred. No. 9.4;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 THNRACRGTG 11
DB 102 TQDTVCRCRAG 112

RESULT 15
FASL_RAT STANDARD; PRT; 324 AA.
ID FASL_RAT
AC Q63199;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE FASL RECEPTOR PRECURSOR (APOPTOSIS-MEDIATING SURFACE ANTIGEN FAS)
DE (APO-1 ANTIGEN) (CD95).
GN TNFRSF6 OR FTL OR FAS.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
[1]
SEQUENCE FROM N.A.
STRAIN=SPRAGUE-DAWLEY; TISSUE=LIVER;
MEDLINE-94128114; PubMed-7507668;
Kimura K., Yamamoto M., Wakatsuki T.;
"A variant mRNA species encoding a truncated form of Fas antigen in
the rat liver."
Biochem. Biophys. Res. Commun. 198:666-674(1994).
RL Biochem. Biophys. Res. Commun. 198:666-674(1994).
CC -!- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS FASL. THE
CC ADAPTOR MOLECULE FADD RECRUITS CASPASE-8 TO THE ACTIVATED
CC RECEPTOR. THE RESULTING AGGREGATE CALLED THE DEATH-INDUCING
CC SIGNALING COMPLEX (DISC) PERFORMS CASPASE-8 PROTEOLYTIC
CC ACTIVATION. ACTIVE CASPASE-8 INITIATES THE SUBSEQUENT CASCADE OF
CC CASPASES (ASPARTATE-SPECIFIC CYSTEINE PROTEASES) MEDIATING
CC APOPTOSIS. FAS-MEDIATED APOPTOSIS MAY HAVE A ROLE IN THE
CC INDUCTION OF PERIPHERAL TOLERANCE. IN THE ANTIGEN-STIMULATED
CC SUICIDE OF MATURE T-CELLS, OR BOTH (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- DOMAIN: CONTAINS A DEATH DOMAIN INVOLVED IN THE BINDING OF FADD,
CC AND MAYBE TO OTHER CYTOSOLIC ADAPTOR PROTEINS.


```

CC  -!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC  -!- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; D26112; BAA05108.1; -
DR  INTERPRO; IPR000488; -
DR  INTERPRO; IPR01368; -
DR  PFAM; PF00020; TNFR_C6; 3.
DR  PFAM; PF00531; death; 1.
DR  PROSITE; PS00652; TNFR_NGFR_1; 2.
DR  PROSITE; PS50050; TNFR_NGFR_2; 2.
DR  PROSITE; PS50017; DEATH_DOMAIN; 1.
KW  Apoptosis; Receptor; Glycoprotein; Transmembrane; Repeat; Signal.
FT  SIGNAL 1 21 BY SIMILARITY.
FT  CHAIN 22 324 FASL RECEPTOR.
FT  DOMAIN 22 171 EXTRACELLULAR (POTENTIAL).
FT  TRANSMEM 172 188 POTENTIAL.
FT  DOMAIN 189 324 CYTOPLASMIC (POTENTIAL).
FT  DOMAIN 43 163 3 X TNFR-CYS.
FT  REPEAT 43 79 TNFR-CYS 1.
FT  REPEAT 80 123 TNFR-CYS 2.
FT  REPEAT 124 163 TNFR-CYS 3.
FT  DOMAIN 219 303 DEATH DOMAIN.
FT  CARBOHYD 43 43 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 132 132 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ  SEQUENCE 324 AA; 36835 MW; D25D583C909D9D09 CRC64;

```

```

Query Match          53.4%; Score 39; DB 1; Length 324;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```

```

Oy  1 THNRACRCRTGF 12
    | | | | |
Db  118 TQNTKCRCKENF 129

```

```

Search completed: January 30, 2001, 17:06:12
Job time: 1233 sec

```

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:55:02 ; Search time 254.27 Seconds

(without alignments)
5.532 Million cell updates/sec

Title: US-09-518-931-4_COPY_108_119

Perfect score: 73

Sequence: 1 THNRACRCRTGF 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_15:

- 1: sp_archaea:
- 2: sp_bacteria:
- 3: sp_fungi:
- 4: sp_human:
- 5: sp_invertebrate:
- 6: sp_mammal:
- 7: sp_mhc:
- 8: sp_organelle:
- 9: sp_phase:
- 10: sp_plant:
- 11: sp_rodent:
- 12: sp_virus:
- 13: sp_vertebrate:
- 14: sp_unclassified:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	73	100.0	300	4 Q95407	Q95407 homo sapien
2	50	68.5	4599	4 Q9NZR2	Q9NZR2 homo sapien
3	50	68.5	4599	11 Q9J118	Q9J118 mus musculu
4	49	67.1	372	4 Q9UHP4	Q9UHP4 homo sapien
5	49	67.1	401	4 Q00300	Q00300 homo sapien
6	48	65.8	326	12 Q57120	Q57120 cowpox viru
7	48	65.8	326	12 Q57122	Q57122 cowpox viru
8	48	65.8	347	12 Q57119	Q57119 cowpox viru
9	48	65.8	348	12 Q57277	Q57277 monkeypox v
10	48	65.8	348	12 Q57103	Q57103 monkeypox v
11	48	65.8	348	12 Q57108	Q57108 monkeypox v
12	48	65.8	348	12 Q57112	Q57112 variola vir
13	48	65.8	348	12 Q85407	Q85407 variola vir
14	48	65.8	349	12 Q57284	Q57284 camelppox vi
15	48	65.8	349	12 Q57291	Q57291 monkeypox v
16	48	65.8	349	12 Q57097	Q57097 camelppox vi
17	48	65.8	349	12 Q57098	Q57098 camelppox vi
18	48	65.8	349	12 Q57099	Q57099 monkeypox v
19	48	65.8	349	12 Q57100	Q57100 monkeypox v

20	48	65.8	349	12	Q57101	Q57101 monkeypox v
21	48	65.8	349	12	Q57102	Q57102 monkeypox v
22	48	65.8	349	12	Q57109	Q57109 variola vir
23	48	65.8	349	12	Q57110	Q57110 variola vir
24	48	65.8	349	12	Q57111	Q57111 variola vir
25	48	65.8	349	12	Q89118	Q89118 variola vir
26	48	65.8	349	12	Q89098	Q89098 variola vir
27	48	65.8	351	12	Q57117	Q57117 cowpox viru
28	48	65.8	351	12	Q57121	Q57121 cowpox viru
29	48	65.8	351	12	Q73559	Q73559 cowpox viru
30	48	65.8	360	12	Q57118	Q57118 cowpox viru
31	47	64.4	401	11	Q08712	Q08712 mus musculu
32	47	64.4	401	11	Q08727	Q08727 rattus norv
33	47	64.4	3767	5	Q9UAL3	Q9UAL3 caenorhabdi
34	46	63.0	347	12	Q57115	Q57115 cowpox viru
35	46	63.0	349	12	Q57305	Q57305 cowpox viru
36	46	63.0	350	12	Q57116	Q57116 cowpox viru
37	46	63.0	350	12	Q57123	Q57123 cowpox viru
38	46	63.0	355	12	Q85308	Q85308 cowpox viru
39	46	63.0	4545	11	Q61291	Q61291 mus musculu
40	45	61.6	527	13	Q9PSV0	Q9PSV0 gallus gall
41	45	61.6	529	13	Q91436	Q91436 gallus gall
42	45	61.6	585	11	Q35675	Q35675 mus musculu
43	45	61.6	685	11	Q9JH27	Q9JH27 mus musculu
44	45	61.6	686	11	Q9J171	Q9J171 mus musculu
45	45	61.6	2451	12	Q9WJD4	Q9WJD4 venezuelan

ALIGNMENTS

RESULT 1

Q95407	Q95407	PRELIMINARY;	PRT;	300 AA.
ID	Q95407			
AC	Q95407;			
DT	01-MAY-1999 (TREMBLrel. 10, Created)			
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)			
DT	01-OCT-2000 (TREMBLrel. 15, Last annotation update)			
DE	DECOY RECEPTOR 3 (M68) (M68C) (M68E).			
GN	DCR3 OR TR6.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
EN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99087326; PubMed=9872321;			
RA	Pitti R.M., Marsters S.A., Lawrence D.A., Roy M., Kischkel F.C.,			
RA	Dowd P., Huang A., Donahue C.J., Sherwood S.W., Baldwin D.T.,			
RA	Godowski P.J., Wood W.I., Gurney A.L., Hillan K.J., Cohen R.L.,			
RA	Goddard A.D., Botstein D., Ashkenazi A.			
RT	"Genomic amplification of a decoy receptor for Fas ligand in lung and			
RL	colon cancer."			
RL	Nature 396:699-703(1998).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Blood.			
RX	MEDLINE=99253915; PubMed=10318773;			
RA	Yu K.Y., Kwon B., Ni J., Zhai Y., Ebner R., Kwon B.S.;			
RT	"A newly identified member of tumor necrosis factor receptor			
RL	superfamily (TR6) suppresses LIGHT-mediated apoptosis."			
RL	J. Biol. Chem. 274:13733-13736(1999).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=PANCREAS;			
RX	MEDLINE=20122600; PubMed=10655513;			
RA	Bai C., Connolly B., Metzker M.L., Hilliard C.A., Liu X., Sandig V.,			
RA	Soderman A., Galloway S.M., Liu O., Austin C.P., Caskey C.T.;			
RT	"Overexpression of M68/Dcr3 in human gastrointestinal tract tumors			
RL	independent of gene amplification and its location in a four-gene			
RL	cluster."			
RL	Proc. Natl. Acad. Sci. U.S.A. 97:1230-1235(2000).			
DR	EMBL; AF104419; AAD03056.1; -.			

DR EMBL; AF134240; AAD29688.1; -;
 DR EMBL; AF217796; AAF35244.1; -;
 DR EMBL; AF217793; AAF33685.1; -;
 DR EMBL; AF217794; AAF33686.1; -;
 DR HSSP; P25942; ICDF;
 DR INTERPRO; IPR000561; -;
 DR INTERPRO; IPR001368; -;
 DR PFAM; PF00020; TNFR_C6; 4;
 DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1;
 DR PROSITE; PS01186; EGF_2; UNKNOWN_1;
 DR PROSITE; PS50050; TNFR_NGFR_2; 2;
 DR PRODOM; PD000771; -; 1;
 KW Receptor.
 SQ SEQUENCE 300 AA; 32679 MW; F90AEE33718449AF CRC64;

Query Match 100.0%; Score 73; DB 4; Length 300;
 Best Local Similarity 100.0%; Pred. No. 6.1e-06;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 THNRACRRTGF 12
 Db 108 THNRACRRTGF 119

RESULT 2
 ID Q9NZR2 PRELIMINARY; PRT; 4599 AA.
 AC Q9NZR2;
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
 DE LOW DENSITY LIPOPROTEIN RECEPTOR RELATED PROTEIN-DELETED IN TUMOR.
 GN LRPDIT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-2027253; PubMed=10766186;
 RA Liu C.X., Musco S., Lisitsyn N.M., Forgacs E., Minna J.D.,
 RA Lisitsyn N.A.;
 RT "LRP-DIT, a putative endocytic receptor gene, is frequently
 RT inactivated in non-small cell lung cancer cell lines.";
 RL Cancer Res. 60:1961-1967(2000).
 DR EMBL; AF176832; AAF70379.1; -;
 KW Receptor; Lipoprotein.
 SQ SEQUENCE 4599 AA; 515504 MW; 9A97A555F1FAAE66 CRC64;

Query Match 68.5%; Score 50; DB 4; Length 4599;
 Best Local Similarity 88.9%; Pred. No. 0.86;
 Matches 8; Conservative 0; Mismatches 0; Indels 1; Gaps 0;

Qy 4 RACRCRTGF 12
 Db 499 RTCRCRTGF 507

RESULT 3
 ID Q9J118 PRELIMINARY; PRT; 4599 AA.
 AC Q9J118;
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
 DE LOW DENSITY LIPOPROTEIN RECEPTOR RELATED PROTEIN LRPIB/LRP-DIT.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]

RP SEQUENCE FROM N.A.
 RA Yaklichkin S., Lisitsyn N.;
 RT "The structure of the mouse homologue of the human candidate tumor
 RT suppressor gene LRPIB/LRP-DIT.";
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF270884; AAF81758.1; -;
 KW Receptor; Lipoprotein.
 SQ SEQUENCE 4599 AA; 513629 MW; 11462A3354FFB200 CRC64;

Query Match 68.5%; Score 50; DB 11; Length 4599;
 Best Local Similarity 88.9%; Pred. No. 0.86;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RACRCRTGF 12
 Db 499 RTCRCRTGF 507

RESULT 4
 ID Q9UHP4 PRELIMINARY; PRT; 372 AA.
 AC Q9UHP4;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
 DE OSTEOPROTEGERIN (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA;
 RA He Z.-Y., Yang G.-Z., Zhang W.-J., Wu X.-F.;
 RT "Cloning and Expression of Osteoprotegerin from Homo sapiens.";
 RL Sheng Wu Hua Hsueh Yu Sheng Wu Wu Li Hsueh pao 31:680-684(1999).
 DR EMBL; AF134187; AAF20168.1; -;
 DR HSSP; P25942; ICDF.
 DR INTERPRO; IPR001368; -;
 DR PFAM; PF00020; TNFR_C6; 3;
 DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
 DR PROSITE; PS50050; TNFR_NGFR_2; 2;
 FT NON_TER 1
 SQ SEQUENCE 372 AA; 42758 MW; F02527A5CD01CCD3 CRC64;

Query Match 67.1%; Score 49; DB 4; Length 372;
 Best Local Similarity 63.6%; Pred. No. 0.19;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 THNRACRRTG 11
 Db 79 THNRVCEKEG 89

RESULT 5
 ID O00300 PRELIMINARY; PRT; 401 AA.
 AC O00300; O60236;
 DT 01-JUL-1997 (TRENBLrel. 04, Created)
 DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)
 DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
 DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)
 DE (OCIF) (TUMOR NECROSIS FACTOR RECEPTOR SUPERFAMILY MEMBER 11b).
 GN TNFRSF11B OR OPG OR OCIF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;

RA MEDLINE=97262071; PubMed=9108485;
RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,
RA Lunethy R., Nguyen H.Q., Wooden S., Bennett L., Boone T., Shimamoto G.,
RA Derose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,
RA Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,
RA Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,
RA Sugis S., Boyle W.J.;
RA "Osteoprotegerin: a novel secreted protein involved in the regulation
RT of bone density";
RL Cell 89:309-319(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG FIBROBLAST;
RX MEDLINE=98151033; PubMed=9492069;
RA Yasuda H., Shima N., Nakagawa N., Mochizuki S.-I., Yano K., Fujise N.,
RA Sato Y., Goto M., Yamauchi K., Kuriyama M., Kanno T., Murakami A.,
RA Tsuda E., Morinaga T., Higashio K.;
RT "Identity of osteoclastogenesis inhibitory factor (OCIF) and
RT osteoprotegerin (OPG): a mechanism by which OPG/OCIF inhibits
RT osteoclastogenesis in vitro";
RL Endocrinology 139:1329-1337(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=98351569; PubMed=9688283;
RA Morinaga T., Nakagawa N., Yasuda H., Tsuda E., Higashio K.;
RT "Cloning and characterization of the gene encoding human
RT osteoprotegerin/osteoclastogenesis-inhibitory factor";
RL Eur. J. Biochem. 254:685-691(1998).
RN [4]
RP FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES
RC OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY
RX SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO
RT OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN
RC STROMAL CELLS AND OSTEOCLAST PROGENITORS.
RX -1- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM) (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGHEST LEVELS IN LUNG, HEART,
CC KIDNEY, PLACENTA, THYROID, SPINAL CORD AND LIVER. ALSO DETECTED IN
CC A NUMBER OF OTHER HEMATOPOIETIC AND IMMUNE ORGANS. NOT DETECTED IN
CC THE PANCREAS OR PERIPHERAL BLOOD LYMPHOCYTES.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL: AB002146; BAA25910.1; -;
DR EMBL: AB008822; BAA32076.1; -;
DR EMBL: AB008821; BAA32076.1; JOINED.
DR EMBL: U94332; AAB53709.1; -;
DR HSSP: P25942; ICDF.
DR MIN: 602643; -;
DR INTERPRO: IPR001368; -;
DR PFAM: PF00020; TNFR_C6; 3.
DR PROSITE: PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE: PS50050; TNFR_NGFR_2; 2.
DR PRODOM: PD000771; -; 1.
KW Glycoprotein; Repeat; Cytokine; Signal.
FT SIGNAL 1 21
FT CHAIN 22 401 BY SIMILARITY.
FT DOMAIN 23 183 OSTEOPROTEGERIN.
FT REPEAT 23 63 4 X TNFR-CYS.
FT REPEAT 64 106 TNFR-CYS 1.
FT REPEAT 107 143 TNFR-CYS 2.
FT REPEAT 144 201 TNFR-CYS 3.
FT REPEAT 306 365 TNFR-CYS 4.
FT DOMAIN 41 54 DEATH DOMAIN.
FT DISULFID 44 62 BY SIMILARITY.
FT DISULFID 65 80 BY SIMILARITY.
FT DISULFID 83 97 BY SIMILARITY.
FT DISULFID 87 105 BY SIMILARITY.
FT DISULFID 118 142 BY SIMILARITY.
FT DISULFID 145 160 BY SIMILARITY.
FT CARBOHYD 98 98 POTENTIAL.
FT CARBOHYD 152 152 POTENTIAL.
FT CARBOHYD 165 165 POTENTIAL.
FT CARBOHYD 178 178 POTENTIAL.

FT CARBOHYD 289 289 POTENTIAL.
FT CONFLICT 263 263 A -> D (IN REF. 2 AND 3).
SQ SEQUENCE 401 AA; 45996 MW; EB42FA51C9D7C71E CRC64;

Query Match 67.1%; Score 49; DB 4; Length 401;
Best Local Similarity 63.6%; Pred. No. 0.2;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 THNRACRGTG 11
| | | | | | | | | |
Db 100 THNRVCECKG 110

RESULT 6
O57120 PRELIMINARY; PRT; 326 AA.
ID O57120 AC O57120;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN CRMB.
OS Cowpox virus (CPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10243;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MUNICH 89/1(CAT);
RA Loparev V.N., Parsons J.M., Esposito J.J.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U90231; AAB94387.1; -;
DR HSSP: P19438; ITNR.
DR INTERPRO: IPR001368; -;
DR PFAM: PF00020; TNFR_C6; 2.
DR PROSITE: PS00652; TNFR_NGFR_1; UNKNOWN_2.
DR PROSITE: PS50050; TNFR_NGFR_2; 1.
DR PRODOM: PD000771; -; 1.
SQ SEQUENCE 326 AA; 35512 MW; 1F800EEDA330F5A5 CRC64;

Query Match 65.8%; Score 48; DB 12; Length 326;
Best Local Similarity 58.3%; Pred. No. 0.26;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 THNRACRGTG 12
| | | | | | | | | |
Db 80 THNRICEGSPG 91

RESULT 7
O57122 PRELIMINARY; PRT; 326 AA.
ID O57122 AC O57122;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN CRMB.
OS Cowpox virus (CPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10243;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MUNICH OPV90/1(CAT);
RA Loparev V.N., Parsons J.M., Esposito J.J.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U90234; AAB94390.1; -;
DR HSSP: P19438; ITNR.
DR INTERPRO: IPR001368; -;
DR PFAM: PF00020; TNFR_C6; 2.

```

DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_2.
DR PROSITE; PS00500; TNFR_NGFR_2; 1.
DR PRODOM; PD000771; -; 1.
SQ SEQUENCE 326 AA; 35567 MW; 1F08BF0C7AB0C145 CRC64;

Query Match 65.8%; Score 48; DB 12; Length 326;
Best Local Similarity 58.3%; Pred. No. 0.26;
Matches 7; Conservative 1; Mismatches 0; Gaps 0;

QY 1 THNRACRCRTGF 12
    |||| | | |
DB 80 THNRICECPGY 91

RESULT 8
O57119 ID O57119 PRELIMINARY; PRT; 347 AA.
AC O57119;
DT 01-JUN-1998 (TRENBLrel. 06, Created)
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN CRMB.
OS Cowpox virus (CPV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10243;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MUNICH EP2-1975(ELEPHANT);
RA Loparev V.N., Parsons J.M., Esposito J.J.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U90230; AAB94386.1; -.
DR HSSP; P19438; 1NCF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 2.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_2.
DR PROSITE; PS00500; TNFR_NGFR_2; 1.
DR PRODOM; PD000771; -; 1.
SQ SEQUENCE 347 AA; 37971 MW; A51B36A5619F48DA CRC64;

Query Match 65.8%; Score 48; DB 12; Length 347;
Best Local Similarity 58.3%; Pred. No. 0.27;
Matches 7; Conservative 1; Mismatches 0; Gaps 0;

QY 1 THNRACRCRTGF 12
    |||| | | |
DB 102 THNRICECPGY 113

RESULT 9
O57277 ID O57277 PRELIMINARY; PRT; 348 AA.
AC O57277;
DT 01-JUN-1998 (TRENBLrel. 06, Created)
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN CRMB.
OS Monkeypox virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10244;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZAIRE-1996(96-17), ZAIRE-1996 (96-16);
RA Loparev V.N., Parsons J.M., Esposito J.J.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; U88543; AAB94378.1; -.
DR EMBL; U87841; AAB94358.1; -.
DR HSSP; P25942; 1CDF.

DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 2.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00500; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -; 1.
SQ SEQUENCE 348 AA; 38212 MW; 54019521556C2D8F CRC64;

Query Match 65.8%; Score 48; DB 12; Length 348;
Best Local Similarity 58.3%; Pred. No. 0.27;
Matches 7; Conservative 1; Mismatches 0; Gaps 0;

QY 1 THNRACRCRTGF 12
    |||| | | |
DB 103 THNRICECPGY 114

RESULT 10
O57103 ID O57103 PRELIMINARY; PRT; 348 AA.
AC O57103;
DT 01-JUN-1998 (TRENBLrel. 06, Created)
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN CRMB.
OS Monkeypox virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10244;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZAIRE-1979 (79-0005);
RA Loparev V.N., Parsons J.M., Esposito J.J.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U87847; AAB94364.1; -.
DR HSSP; P25942; 1CDF.
DR INTERPRO; IPR001368; -.
DR PFAM; PF00020; TNFR_C6; 2.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00500; TNFR_NGFR_2; 2.
DR PRODOM; PD000771; -; 1.
SQ SEQUENCE 348 AA; 38184 MW; 34A5E668B27907B5 CRC64;

Query Match 65.8%; Score 48; DB 12; Length 348;
Best Local Similarity 58.3%; Pred. No. 0.27;
Matches 7; Conservative 1; Mismatches 0; Gaps 0;

QY 1 THNRACRCRTGF 12
    |||| | | |
DB 103 THNRICECPGY 114

RESULT 11
O57108 ID O57108 PRELIMINARY; PRT; 348 AA.
AC O57108;
DT 01-JUN-1998 (TRENBLrel. 06, Created)
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN CRMB.
OS Monkeypox virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OX NCBI_TaxID=10244;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ZAIRE-1970 (CONGO-8);
RA Loparev V.N., Parsons J.M., Esposito J.J.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U88142; AAB94367.1; -.

```

DR HSSP; P25942; 1CDF.
 DR INTERPRO; IPR001368; -.
 DR PFAM; PF00020; TNFR_C6; 2.
 DR PROSITE; PS00652; TNFR_NGFR_1; 2.
 DR PROSITE; PS50050; TNFR_NGFR_2; 2.
 DR PRODOM; PD000771; -. 1.
 SQ SEQUENCE 348 AA; 38212 MW; B555979057DEC91F CRC64;

Query Match 65.8%; Score 48; DB 12; Length 348;
 Best Local Similarity 58.3%; Pred. No. 0.27;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 THNRACRCRTGF 12
 ||||| | | | |
 Db 103 THNRICECPGY 114

RESULT 12

ID O57112 PRELIMINARY; PRT; 348 AA.
 AC O57112;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
 GN CRMB.
 OS Variola virus.
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Orthopoxvirus.
 OX NCBI_TaxID=10255;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN='WHITEPOX' CHIMP 9-2 (CH9-2);
 RA Loparev V.N., Parsons J.M., Esposito J.J.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U88151; AAB94376.1; -.
 DR HSSP; P25942; 1CDF.
 DR INTERPRO; IPR001368; -.
 DR PFAM; PF00020; TNFR_C6; 2.
 DR PROSITE; PS00652; TNFR_NGFR_1; 2.
 DR PROSITE; PS50050; TNFR_NGFR_2; 2.
 DR PRODOM; PD000771; -. 1.
 SQ SEQUENCE 348 AA; 38041 MW; 92D333794FC373E3 CRC64;

Query Match 65.8%; Score 48; DB 12; Length 348;
 Best Local Similarity 58.3%; Pred. No. 0.27;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 THNRACRCRTGF 12
 ||||| | | | |
 Db 102 THNRICECPGY 113

RESULT 13

ID Q85407 PRELIMINARY; PRT; 348 AA.
 AC Q85407;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE HOMOLOG OF VACCINIA VIRUS CDS B28R.
 GN G2R OR CRMB.
 OS Variola virus.
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Orthopoxvirus.
 OX NCBI_TaxID=10255;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BANGLADESH-1975;
 RX MEDLINE=94088747; PubMed=8264798;
 RA Massung R.F., Esposito J.J., Liu L.I., Qi J., Utterback T.R.,

RA Knight J.C., Aubin L., Yuran T.E., Parsons J.M., Loparev V.N.;
 RT "Potential virulence determinants in terminal regions of variola
 RT smallpox virus genome";
 RL Nature 366:748-751(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN='WHITEPOX' CHIMP 9-4 (CH9-4);
 RA Loparev V.N., Parsons J.M., Esposito J.J.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; L22579; AAA60933.1; -.
 DR EMBL; U88152; AAB94377.1; -.
 DR HSSP; P25942; 1CDF.
 DR INTERPRO; IPR001368; -.
 DR PFAM; PF00020; TNFR_C6; 2.
 DR PROSITE; PS00652; TNFR_NGFR_1; 2.
 DR PROSITE; PS50050; TNFR_NGFR_2; 2.
 DR PRODOM; PD000771; -. 1.
 SQ SEQUENCE 348 AA; 38075 MW; A9BE85AF94AB5369 CRC64;

Query Match 65.8%; Score 48; DB 12; Length 348;
 Best Local Similarity 58.3%; Pred. No. 0.27;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 THNRACRCRTGF 12
 ||||| | | | |
 Db 102 THNRICECPGY 113

RESULT 14

ID O57284 PRELIMINARY; PRT; 349 AA.
 AC O57284;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
 GN CRMB.
 OS Camelpox virus (strain CP-1).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Orthopoxvirus.
 OX NCBI_TaxID=28873;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DUBAI-1992 (CP-5), SOMALIA-1978;
 RA Loparev V.N., Parsons J.M., Esposito J.J.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U87840; AAB94357.1; -.
 DR EMBL; U87837; AAB94354.1; -.
 DR HSSP; P25942; 1CDF.
 DR INTERPRO; IPR001368; -.
 DR PFAM; PF00020; TNFR_C6; 2.
 DR PROSITE; PS00652; TNFR_NGFR_1; 2.
 DR PROSITE; PS50050; TNFR_NGFR_2; 2.
 DR PRODOM; PD000771; -. 1.
 SQ SEQUENCE 349 AA; 38036 MW; EA412AEE86E090E4 CRC64;

Query Match 65.8%; Score 48; DB 12; Length 349;
 Best Local Similarity 58.3%; Pred. No. 0.27;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 THNRACRCRTGF 12
 ||||| | | | |
 Db 103 THNRICECPGY 114

RESULT 15

ID O57291 PRELIMINARY; PRT; 349 AA.
 AC O57291;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)

DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
 GN CRMB.
 OS Monkeypox virus.
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Orthopoxvirus.
 OX NCBI_TaxID=10244;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VARIOUS STRAINS;
 RA Loparev V.N., Parsons J.M., Esposito J.J.;
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U88144; AAB94369.1; -
 DR EMBL; U87842; AAB94359.1; -
 DR EMBL; U87994; AAB94365.1; -
 DR EMBL; U87995; AAB94366.1; -
 DR EMBL; U88143; AAB94368.1; -
 DR HSSP; P25942; ICDF.
 DR INTERPRO; IPR001368; -
 DR PFAM; PF00020; TNFR_G6; 2.
 DR PROSITE; PS00652; TNFR_NGFR_1; 2.
 DR PROSITE; PS00050; TNFR_NGFR_2; 2.
 DR PRODOM; PD000771; -; 1
 SQ SEQUENCE 349 AA; 38295 MW; CBD2C949ED2B8E7C CRC64;

Query Match 65.8%; Score 48; DB 12; Length 349;
 Best Local Similarity 58.3%; Pred. No. 0.27;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 THNRACRCRTGF 12
 DB 103 THNRICECPGY 114

Search completed: January 30, 2001, 16:55:02
 Job time: 918 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:47:50 ; Search time 132.77 seconds
(without alignments)
1.623 Million cell updates/sec

Title: US-09-518-931-4_COPY_108_119

Perfect score: 73

Sequence: 1 THNRACRCRTGF 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued_Patents_AA:*
- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
 - 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
 - 3: /cgn2_6/ptodata/2/iaa/6_COMB.pep.*
 - 4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
 - 5: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	73	100.0	300	2	US-08-794-796-2
2	49	67.1	401	3	US-08-974-022-6
3	47	64.4	207	3	US-08-974-022-47
4	47	64.4	325	1	US-08-292-549-2
5	47	64.4	325	4	PCT-US91-02207-2
6	47	64.4	401	3	US-08-974-022-2
7	47	64.4	401	3	US-08-974-022-4
8	46	63.0	355	1	US-08-292-549-6
9	46	63.0	4544	1	US-08-469-486-52
10	46	63.0	4544	2	US-08-469-658-52
11	44	60.3	685	3	US-08-872-855-2
12	43	58.9	191	3	US-08-974-022-52
13	43	58.9	256	1	US-08-236-318A-6
14	43	58.9	256	4	PCT-US96-03965-2
15	41	56.2	713	3	US-08-872-855-5
16	40	54.8	44	2	US-08-484-434C-24
17	40	54.8	88	1	US-08-469-202-20
18	39	53.4	119	2	US-08-219-237B-3
19	39	53.4	219	3	US-08-974-022-45
20	39	53.4	277	2	US-08-147-784-2
21	39	53.4	314	1	US-08-444-231-19
22	39	53.4	314	1	US-08-152-443A-19
23	39	53.4	314	4	PCT-US95-17083-4
24	39	53.4	335	2	US-08-219-237B-2
25	39	53.4	335	2	US-08-409-338-1
26	39	53.4	335	3	US-08-815-469-6
27	39	53.4	335	4	PCT-US95-17083-2
28	38	52.1	120	3	US-08-974-022-42

29	38	52.1	132	1	US-08-208-008C-2	Sequence 2, Appli
30	38	52.1	163	2	US-08-219-237B-5	Sequence 5, Appli
31	38	52.1	164	2	US-08-232-087A-9	Sequence 9, Appli
32	38	52.1	206	1	US-08-097-827-7	Sequence 7, Appli
33	38	52.1	206	1	US-08-494-574-7	Sequence 7, Appli
34	38	52.1	227	3	US-08-974-022-48	Sequence 48, Appli
35	38	52.1	241	2	US-08-460-309-17	Sequence 17, Appli
36	38	52.1	241	2	US-08-125-077-17	Sequence 17, Appli
37	38	52.1	380	3	US-08-468-846-2	Sequence 2, Appli
38	38	52.1	388	1	US-08-290-448A-80	Sequence 80, Appli
39	38	52.1	388	1	US-08-290-448A-80	Sequence 80, Appli
40	38	52.1	388	1	US-08-175-069A-80	Sequence 80, Appli
41	38	52.1	438	1	US-08-097-827-11	Sequence 11, Appli
42	38	52.1	438	1	US-08-494-574-11	Sequence 11, Appli
43	38	52.1	454	3	US-08-276-968A-22	Sequence 22, Appli
44	38	52.1	461	1	US-08-385-229-2	Sequence 2, Appli
45	38	52.1	461	5	5395760-2	Patent No. 5395760

ALIGNMENTS

RESULT 1
US-08-794-796-2
; Sequence 2, Application US/08794796
; Patent No. 5885800
; GENERAL INFORMATION:
; APPLICANT: Emery, John
; APPLICANT: Tan, KB
; APPLICANT: Truneh, Alem
; APPLICANT: Young, Peter
; TITLE OF INVENTION: Tumor Necrosis Related Receptor,
; TITLE OF INVENTION: TR4
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/794,796
FILING DATE: 04-FEB-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Han, William T
REGISTRATION NUMBER: 34,344
REFERENCE/DOCKET NUMBER: GH50000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5219
TELEFAX: 610-270-4026
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 300 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-794-796-2

Query Match 100.0%; Score 73; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 0.00034;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 THNRACRTG 12
| | | | | | | | | |
DB 108 THNRACRTG 119

RESULT 2

US-08-974-022-6
; Sequence 6, Application US/08974022
; Patent No. 6015938

; GENERAL INFORMATION:
; APPLICANT: Boyle, William J.
; APPLICANT: Lacey, David L.
; APPLICANT: Calzone, Frank J.
; APPLICANT: Chang, Ming-Shi
; TITLE OF INVENTION: OSTEOPROTEGERIN
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 Dehaviiland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA

; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,022
; FILING DATE: 12-DEC-1995

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/577,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-378
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 401 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-974-022-6

Query Match 67.1%; Score 49; DB 3; Length 401;
Best Local Similarity 63.6%; Pred. No. 1.7;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 THNRACRTG 11
| | | | | | | | | |
DB 100 THNRVCECKG 110

RESULT 3

US-08-974-022-47
; Sequence 47, Application US/08974022
; Patent No. 6015938

; GENERAL INFORMATION:
; APPLICANT: Boyle, William J.
; APPLICANT: Lacey, David L.
; APPLICANT: Calzone, Frank J.
; APPLICANT: Chang, Ming-Shi
; TITLE OF INVENTION: OSTEOPROTEGERIN
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 Dehaviiland Drive
; CITY: Thousand Oaks

; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,022
; FILING DATE: 12-DEC-1995

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/577,788
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-378
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 207 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-974-022-47

Query Match 64.4%; Score 47; DB 3; Length 207;
Best Local Similarity 63.6%; Pred. No. 1.9;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 THNRACRTG 11
| | | | | | | | | |
DB 99 THDRVCNCSTG 109

RESULT 4

US-08-292-549-2
; Sequence 2, Application US/08292549
; Patent No. 5464938

; GENERAL INFORMATION:
; APPLICANT: Smith, Craig A.
; APPLICANT: Goodwin, Raymond G.
; TITLE OF INVENTION: Isolated Viral Protein TNF Antagonists
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/292,549
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/963,330
; FILING DATE: 10/19/92
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia A.
; REGISTRATION NUMBER: 34,693
; REFERENCE/DOCKET NUMBER: 2602-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 325 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-292-549-2

Query Match 64.4%; Score 47; DB 1; Length 325;
Best Local Similarity 63.6%; Pred. No. 2.9;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 THNRACRGTG 11
||| | | | |
Db 99 THDRVCNCSTG 109

RESULT 5
PCT-US91-02207-2
Sequence 2, Application PC/TUS9102207
GENERAL INFORMATION:
APPLICANT: Smith, Craig A.
APPLICANT: Goodwin, Raymond G.
TITLE OF INVENTION: Isolated Viral Protein Cytokine Antagonists
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98101

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/02207
FILING DATE: 19910329
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Wight, Christopher L.
REGISTRATION NUMBER: 31,680
REFERENCE/DOCKET NUMBER: 2602
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 587-0606
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 325 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US91-02207-2

Query Match 64.4%; Score 47; DB 4; Length 325;
Best Local Similarity 63.6%; Pred. No. 2.9;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 THNRACRGTG 11
||| | | | |
Db 99 THDRVCNCSTG 109

RESULT 6
US-08-974-022-2
Sequence 2, Application US/08974022
GENERAL INFORMATION:
APPLICANT: Boyle, William J.
APPLICANT: Lacey, David L.

APPLICANT: Calzone, Frank J.
APPLICANT: Chang, Ming-Shi
TITLE OF INVENTION: OSTEOPROTEGERIN
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,022
FILING DATE: 12-DEC-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/577,788
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-378
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 401 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-974-022-2

Query Match 64.4%; Score 47; DB 3; Length 401;
Best Local Similarity 63.6%; Pred. No. 3.5;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 THNRACRGTG 11
||| | | | |
Db 100 THNRVCEEG 110

RESULT 7
US-08-974-022-4
Sequence 4, Application US/08974022
Patent No. 6015938
GENERAL INFORMATION:
APPLICANT: Boyle, William J.
APPLICANT: Lacey, David L.
APPLICANT: Calzone, Frank J.
APPLICANT: Chang, Ming-Shi
TITLE OF INVENTION: OSTEOPROTEGERIN
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,022
FILING DATE: 12-DEC-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/577,788

us-09-518-931-4_copy_108_119.ra

Wed Jan 31 13:44:48 2001

```

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-378
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 401 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-974-022-4

Query Match 64.4%; Score 47; DB 3; Length 401;
Best Local Similarity 63.6%; Pred. NO. 3.5;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 THNRACRTG 11
Db 100 THNRVCEEG 110

RESULT 8
US-08-292-549-6
; Sequence 6, Application US/08292549
; Patent No. 5464938
; GENERAL INFORMATION:
; APPLICANT: Smith, Craig A.
; APPLICANT: Goodwin, Raymond G.
; TITLE OF INVENTION: Isolated Viral Protein TNF Antagonists
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/292,549
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION NUMBER: 07/963,330
; FILING DATE: 10/19/92
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia A.
; REGISTRATION NUMBER: 34,693
; REFERENCE/DOCKET NUMBER: 2602-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-292-549-6

Query Match 63.0%; Score 46; DB 1; Length 355;
Best Local Similarity 58.3%; Pred. NO. 4.4;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 THNRACRTG 12
Db 103 THNRICDAPG 114

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-378
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 401 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-974-022-4

Query Match 64.4%; Score 47; DB 3; Length 401;
Best Local Similarity 63.6%; Pred. NO. 3.5;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 THNRACRTG 11
Db 100 THNRVCEEG 110

RESULT 8
US-08-292-549-6
; Sequence 6, Application US/08292549
; Patent No. 5464938
; GENERAL INFORMATION:
; APPLICANT: Smith, Craig A.
; APPLICANT: Goodwin, Raymond G.
; TITLE OF INVENTION: Isolated Viral Protein TNF Antagonists
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/292,549
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION NUMBER: 07/963,330
; FILING DATE: 10/19/92
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia A.
; REGISTRATION NUMBER: 34,693
; REFERENCE/DOCKET NUMBER: 2602-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 587-0430
; TELEFAX: (206) 233-0644
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 355 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-292-549-6

Query Match 63.0%; Score 46; DB 1; Length 355;
Best Local Similarity 58.3%; Pred. NO. 4.4;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 THNRACRTG 12
Db 103 THNRICDAPG 114

RESULT 9
US-08-469-486-52
; Sequence 52, Application US/08469486
; Patent No. 5739281
; GENERAL INFORMATION:
; APPLICANT: Thøgersen, Hans Christian
; APPLICANT: Holtet, Thor Las
; APPLICANT: Etzerodt, Michael
; TITLE OF INVENTION: Improved method for the refolding of
; TITLE OF INVENTION: proteins
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,486
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/192,060
; FILING DATE: February 4, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul T. Clark
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06363/002001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 542 5070
; TELEFAX: 617 542 8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4544 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-469-486-52

Query Match 63.0%; Score 46; DB 1; Length 4544;
Best Local Similarity 77.8%; Pred. NO. 45;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 RACRCRTG 12
Db 502 RTCRCRSGF 510

RESULT 10
US-08-469-658-52
; Sequence 52, Application US/08469658
; Patent No. 5917018
; GENERAL INFORMATION:
; APPLICANT: Thøgersen, Hans Christian
; APPLICANT: Holtet, Thor Las
; APPLICANT: Etzerodt, Michael
; TITLE OF INVENTION: IMPROVED METHOD FOR THE REFOOLDING OF
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
```

STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version
SOFTWARE: #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,658
FILING DATE: June 5, 1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/192,060
FILING DATE: February 4, 1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Paul T. Clark
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 06363/002002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617 542 5070
TELEFAX: 617 542 8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 4544 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-469-658-52

Query Match 63.0%; Score 46; DB 2; Length 4544;
Best Local Similarity 77.8%; Pred. No. 45;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 RACRCRTGF 12
Db 502 RTCRCRSGF 510

RESULT 11
US-08-872-855-2
Sequence 2, Application US/08872855
Patent No. 6121045
GENERAL INFORMATION:
APPLICANT: McCarthy, Sean
APPLICANT: Gearing, David
TITLE OF INVENTION: NOVEL HUMAN DELTA3 COMPOSITIONS AND
TITLE OF INVENTION: THERAPEUTIC USES THEREFOR
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, ROAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/872,855
FILING DATE: 11-JUN-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:

NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: MAA-003.02
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 685 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-872-855-2

Query Match 50.3%; Score 44; DB 3; Length 685;
Best Local Similarity 70.0%; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 NRACRCRTGF 12
Db 423 SRMCRCPGF 432

RESULT 12
US-08-974-022-52
Sequence 52, Application US/08974022
Patent No. 6015938
GENERAL INFORMATION:
APPLICANT: Boyle, William J.
APPLICANT: Lacey, David L.
APPLICANT: Calzone, Frank J.
APPLICANT: Chang, Ming-Shi
TITLE OF INVENTION: OSTEOPROTEGERIN
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,022
FILING DATE: 12-DEC-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/577,788
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-378
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-974-022-52

Query Match 58.9%; Score 43; DB 3; Length 191;
Best Local Similarity 58.3%; Pred. No. 7;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 THNRACRCRTGF 12
Db 1111111111

Db 80 THNAECEIEGF 91

RESULT 13

US-08-236-918A-6
; Sequence 6, Application US/08236918A
; Patent No. 5674704
; GENERAL INFORMATION:
; APPLICANT: Alderson, Mark R.
; APPLICANT: Goodwin, Raymond G.
; APPLICANT: Smith, Craig A.
; TITLE OF INVENTION: Cytokine Designated 4-lBB Ligand
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kathryn A. Anderson, Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98101

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple 7.5.3
SOFTWARE: Microsoft Word, Version #6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/236,918A
FILING DATE: 06-May-1994
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/060,843
FILING DATE: 07-May-1993
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Anderson, Kathryn A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2801-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 256 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-236-918A-6

Query Match 58.9%; Score 43; DB 1; Length 256;
Best Local Similarity 58.3%; Pred. No. 9.1;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 THNRACRRTGF 12

Db 80 THNAECEIEGF 91

RESULT 14

PCT-US96-03965-2
; Sequence 2, Application PC/TUS9603965
; GENERAL INFORMATION:
; APPLICANT: Kwon, Byoung Se
; APPLICANT: Kang, Chang-Yuil
; TITLE OF INVENTION: Monoclonal antibody against human
; TITLE OF INVENTION: receptor 4-lBB
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Barnard, Brown & Michaels
; STREET: 306 East State Street, Suite 220
; CITY: Ithaca
; STATE: NY
; COUNTRY: USA

ZIP: 14850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/03965
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/122,796
FILING DATE: 16-SEP-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/012,269
FILING DATE: 01-FEB-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/922,996
FILING DATE: 30-JUL-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/267,577
FILING DATE: 07-NOV-1988

ATTORNEY/AGENT INFORMATION:
NAME: Michaels, Christopher A.
REGISTRATION NUMBER: 34,390
REFERENCE/DOCKET NUMBER: KW05

TELECOMMUNICATION INFORMATION:
TELEPHONE: 607-273-1711
TELEFAX: 607-273-2609

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 256 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-03965-2

Query Match 58.9%; Score 43; DB 4; Length 256;
Best Local Similarity 58.3%; Pred. No. 9.1;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 THNRACRRTGF 12

Db 80 THNAECEIEGF 91

RESULT 15

US-08-872-855-5
; Sequence 5, Application US/08872855
; Patent No. 6121045
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean
; APPLICANT: Gearing, David
; TITLE OF INVENTION: NOVEL HUMAN DELTA3 COMPOSITIONS AND
; TITLE OF INVENTION: THERAPEUTIC USES THEREFOR
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/872,855
FILING DATE: 11-JUN-1997
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: MAA-003.02
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 713 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-872-855-5

Query Match 56.2%; Score 41; DB 3; Length 713;
Best Local Similarity 85.7%; Pred. No. 46;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 6 CRCRTGF 12
| | | | |
Db 427 CRCOTGF 433

Search completed: January 30, 2001, 16:47:51
Job time: 627 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:45:30 ; Search time 183.67 Seconds
(without alignments)
1.862 Million cell updates/sec

Title: US-09-518-931-4_COPY_129_138

Perfect score: 59

Sequence: 1 HASCPPGAGV 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_36:*

- 1: /SIDS1/gcgdata/geneseq/geneseq/AA1980.DAT:*
- 2: /SIDS1/gcgdata/geneseq/geneseq/AA1981.DAT:*
- 3: /SIDS1/gcgdata/geneseq/geneseq/AA1982.DAT:*
- 4: /SIDS1/gcgdata/geneseq/geneseq/AA1983.DAT:*
- 5: /SIDS1/gcgdata/geneseq/geneseq/AA1984.DAT:*
- 6: /SIDS1/gcgdata/geneseq/geneseq/AA1985.DAT:*
- 7: /SIDS1/gcgdata/geneseq/geneseq/AA1986.DAT:*
- 8: /SIDS1/gcgdata/geneseq/geneseq/AA1987.DAT:*
- 9: /SIDS1/gcgdata/geneseq/geneseq/AA1988.DAT:*
- 10: /SIDS1/gcgdata/geneseq/geneseq/AA1989.DAT:*
- 11: /SIDS1/gcgdata/geneseq/geneseq/AA1990.DAT:*
- 12: /SIDS1/gcgdata/geneseq/geneseq/AA1991.DAT:*
- 13: /SIDS1/gcgdata/geneseq/geneseq/AA1992.DAT:*
- 14: /SIDS1/gcgdata/geneseq/geneseq/AA1993.DAT:*
- 15: /SIDS1/gcgdata/geneseq/geneseq/AA1994.DAT:*
- 16: /SIDS1/gcgdata/geneseq/geneseq/AA1995.DAT:*
- 17: /SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT:*
- 18: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT:*
- 19: /SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT:*
- 20: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT:*
- 21: /SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59	100.0	153	20	Y22222
2	59	100.0	170	19	W63623
3	59	100.0	215	20	W93585
4	59	100.0	245	20	Y28449
5	59	100.0	271	20	Y42184
6	59	100.0	273	20	Y42185
7	59	100.0	300	19	W66102
8	59	100.0	300	19	W63622
9	59	100.0	300	20	Y03099
10	59	100.0	300	20	Y42182
11	59	100.0	300	20	Y17479
12	59	100.0	300	20	Y06817

13	59	100.0	300	20	W97749	Human tumour necro
14	59	100.0	300	20	W95082	Orphan receptor (H
15	59	100.0	300	21	Y77458	Human TNF receptor
16	59	100.0	302	20	Y42183	Human FLINT #2 pro
17	51	86.4	401	18	W38344	Mouse osteoprotegeri
18	49	83.1	401	18	W38343	Rat osteoprotegeri
19	48	81.4	143	17	R99946	Mutated OCIF, OCIF
20	48	81.4	145	17	R99930	Osteoclastogenesis
21	48	81.4	154	17	R99929	Osteoclastogenesis
22	48	81.4	187	17	R99950	Mutated OCIF, OCIF
23	48	81.4	187	21	Y77464	Primate protein se
24	48	81.4	197	17	R99945	Mutated OCIF, OCIF
25	48	81.4	272	17	R99944	Mutated OCIF, OCIF
26	48	81.4	321	17	R99949	Mutated OCIF, OCIF
27	48	81.4	326	17	R99940	Mutated OCIF, OCIF
28	48	81.4	327	17	R99941	Mutated OCIF, OCIF
29	48	81.4	349	20	W83928	Human FTHMA-070 pa
30	48	81.4	351	17	R99943	Mutated OCIF, OCIF
31	48	81.4	359	17	R99937	Mutated OCIF, OCIF
32	48	81.4	359	17	R99939	Mutated OCIF, OCIF
33	48	81.4	360	17	R99936	Mutated OCIF, OCIF
34	48	81.4	380	17	R99924	Mature osteoclasto
35	48	81.4	390	17	R99357	Human tumour necro
36	48	81.4	391	19	W53238	Human OCIF genome
37	48	81.4	393	17	R99948	Mutated OCIF, OCIF
38	48	81.4	395	19	W57636	Modified Trl recep
39	48	81.4	399	17	R99942	Mutated OCIF, OCIF
40	48	81.4	401	17	R99932	Mutated OCIF, OCIF
41	48	81.4	401	17	R99933	Mutated OCIF, OCIF
42	48	81.4	401	17	R99934	Mutated OCIF, OCIF
43	48	81.4	401	17	R99935	Mutated OCIF, OCIF
44	48	81.4	401	17	R99925	Full length osteoc
45	48	81.4	401	20	W83926	Human FTHMA-070 pr

ALIGNMENTS

RESULT 1

Y22222

ID Y22222 standard; Protein; 153 AA.

XX

AC Y22222;

XX

DT 16-SEP-1999 (first entry)

XX

DE Human TNFR superfamily soluble receptor protein sequence.

XX

DE

XX

KW

KW

KW

KW

KW

KW

KW

OS

XX

XX

PN

XX

XX

PD

XX

XX

PF

XX

XX

PR

XX

XX

PA

XX

XX

PI

XX

DR WPI; 1999-405508/34.
 DR N-PSDB; X84621.
 XX
 XX New tumour necrosis factor ligands, useful for induction of cell
 PT death and/or proliferation of cells
 XX
 XX
 PS Claim 1; Page 61; 69pp; English.
 XX
 XX This sequence represents a tumour necrosis factor receptor (TNFR)
 CC superfamily soluble protein of the invention. The invention also relates
 CC to tumour necrosis factor (TNF) ligand (TNFL) family proteins. The TNFL
 CC proteins play regulatory roles in cell proliferation and/or
 CC differentiation, e.g. they can induce production of cytokines,
 CC immunoglobulins, etc. A variety of diseases can be treated by modulating
 CC the activity of TNFL proteins, e.g. they can induce apoptosis of
 CC activated T cells but rescue resting T cell from apoptosis. TNFL
 CC polypeptides can therefore be used to treat autoimmune diseases, such as
 CC myasthenia gravis, insulin-dependent diabetes mellitus, rheumatoid
 CC arthritis, multiple sclerosis, and systemic lupus erythematosus. TNFL
 CC proteins also have tumour stimulating properties, so tumours can be
 CC treated by inhibiting the expression or activity of TNFL. Other
 CC proliferative disorders, such as neoplasias, dysplasias, and hyperplasia
 CC can also be treated using TNFL inhibitors. The TNFL polypeptides and
 CC polynucleotides can also be used to enhance or decrease TNF activity,
 CC thus providing therapeutic benefits such as induction of cell death,
 CC lymphoid organogenesis, or host bacterial resistance, and inhibition of
 CC endotoxin shock, contact hypersensitivity, delayed type sensitivity or
 CC immunocompetence of a transplant recipient. TNF and its receptors play a
 CC major role in host defence and immunosurveillance. As such, there is a
 CC need to identify new members of TNFR families. This invention provides
 CC this need.
 XX
 XX Sequence 153 AA;
 SQ
 Query Match 100.0%; Score 59; DB 20; Length 153;
 Best Local Similarity 100.0%; Pred. No. 0.037;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HASCPPGAGV 10
 Db ||||||||
 45 hascppgagv 54
 RESULT 2
 W63623
 ID W63623 standard; Protein; 170 AA.
 AC
 AC W63623;
 XX
 XX 26-OCT-1998 (first entry)
 DT
 XX Human tumour necrosis factor receptor-6 beta protein.
 DE
 XX Human tumour necrosis factor receptor-6 alpha; TNFR-6 alpha; TNFR-6 beta;
 XX endothelial cells; keratinocytes; normal prostate; apoptosis;
 KW prostate tumour tissue.
 KW
 XX Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 FH Peptide 1..30
 FT Protein 31..170
 FT note= "TNFR-6 beta"
 FT Region 31..166
 FT /note= "Soluble extracellular domain"
 XX
 XX W09830694-A2.
 PN
 XX 16-JUL-1998.
 PD
 XX 13-JAN-1998; 98WO-US00153.
 PF
 XX

PR 14-JAN-1997; 97US-0035496.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Ebner R, Feng P, Gentz RL, Ni J, Ruben SM, Yu G;
 PI
 XX WPI; 1998-399142/34.
 DR N-PSDB; V39086.
 DR
 XX Human tumour necrosis factor receptors 6-alpha and 6-beta - used in
 PT the diagnosis of immune system-related disorder(s)
 PT
 XX Claim 20; Fig 2A-2B; 91pp; English.
 PS
 XX The present sequence represents the human tumour necrosis factor
 CC receptor-6 beta (TNFR-6 beta) protein. The invention also provides
 CC for the TNFR-6 alpha protein (W63622). TNFR-6 alpha and TNFR-6 beta are
 CC members of the tumour necrosis factor receptor (TNFR) family. TNFRs
 CC are expressed in endothelial cells, keratinocytes, normal prostate and
 CC prostate tumour tissue. For a number of disorders of these cells,
 CC particularly of the immune system, substantially altered (whether
 CC increased or decreased) levels of TNFR-6 alpha and/or TNFR-6 beta gene
 CC expression can be detected, therefore the TNFR-6 alpha and TNFR-6 beta
 CC polypeptides, nucleic acids and antibodies are claimed to be useful in
 CC the diagnosis of such disorders. Mutations of the TNFR-6 alpha and
 CC TNFR-6 beta genes can also be detected. The TNFR polypeptides are
 CC also claimed to be useful for identifying ligands which may be useful
 CC in the treatment of apoptosis related disorders.
 XX
 XX Sequence 170 AA;
 SQ
 Query Match 100.0%; Score 59; DB 19; Length 170;
 Best Local Similarity 100.0%; Pred. No. 0.041;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HASCPPGAGV 10
 Db ||||||||
 129 hascppgagv 138
 RESULT 3
 W93585
 ID W93585 standard; Protein; 215 AA.
 AC
 AC W93585;
 XX
 XX 18-JUN-1999 (first entry)
 DT
 XX Human hAPO6 protein.
 DE
 XX Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
 KW developmental abnormality; gestational abnormality; prostate cancer;
 KW APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
 KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
 KW apoptosis; human.
 XX
 XX Homo sapiens.
 OS
 XX W09911791-A2.
 PN
 XX 11-MAR-1999.
 PD
 XX 04-SEP-1998; 98WO-US18393.
 PF
 XX 05-SEP-1997; 97US-0924634.
 PR
 XX (UNIW) UNIV WASHINGTON.
 PA
 XX Chaudhary PM;
 PI
 XX WPI; 1999-205191/17.
 DR N-PSDB; X23419.
 DR

XX New Tumor Necrosis Factor family, receptor polypeptides and ligands -
 PT useful for diagnosis and treatment of prostate cancer and
 PT developmental or gestational abnormalities
 XX
 PS Claim 29; Fig 9; 156pp; English.
 XX
 CC This invention describes isolated Tumor Necrosis Factor (TNF) family
 CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
 CC their active fragments. APO4 is useful for diagnosing prostate cancer
 CC by determining levels of APO4 in an individual. Prostate cancer can also
 CC be treated using APO4 selective binding agents linked to a therapeutic
 CC moiety. APO4 polypeptides are also useful for identifying selective
 CC binding agents, useful in diagnosis/treatment of disease by binding of
 CC agents to the polypeptide/active fragment which is extracellular, or
 CC expressed on the cell surface. The binding is preferably performed in
 CC vivo. APO4 polypeptides/ active fragments are also useful for screening
 CC for agonists and antagonists by binding and observing the change in APO4
 CC activity. Effective pharmacological agents useful in diagnosis or
 CC treatment of disease are also identified using APO4 polypeptides/active
 CC fragments and APO4 signal transducer molecules that specifically interact
 CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4
 CC activity. The method is performed in vivo or in vitro. APO polypeptides
 CC are all useful as immunogens for preparing antibodies. APO4 is also
 CC useful for diagnosis/treatment of developmental or gestational
 CC abnormalities. APO8 was transfected to human breast carcinoma cell line
 CC MCF-7, and induced apoptosis.
 XX
 SQ Sequence 215 AA;

Query Match 100.0%; Score 59; DB 20; Length 215;
 Best Local Similarity 100.0%; Pred. No. 0.051;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HASCPPGAGV 10
 Db 44 hascppgagv 53
 |||||

RESULT 4
 Y28449
 ID Y28449 standard; Protein; 245 AA.
 XX
 AC Y28449;

29-SEP-1999 (first entry)

A human tumour necrosis factor-R2-like proteins (TR2P)-1.

Human tumour necrosis factor-R2-like protein; TR2P; achondroplasia;
 osteoporosis; developmental disorder; Cushing's syndrome;
 muscular dystrophy; epilepsy; hereditary neuropathy;
 Charcot-Marie-Tooth disease; neurofibromatosis; hypothyroidism;
 hydrocephalus; seizure disorder; cerebral palsy; spinal bifida;
 congenital glaucoma; cataract; sensorineural hearing loss;
 reproductive disorder; infertility; ovulatory defect; endometriosis;
 autoimmune disorder; ectopic pregnancy; teratogenesis; spermatogenesis;
 immunological disorder; AIDS; Addison's disease; allergy; bronchitis;
 atherosclerosis; diabetes mellitus; Chron's disease; lupus;
 irritable bowel syndrome; multiple sclerosis; infection;
 neoplastic disorder; adenocarcinoma; leukaemia; lymphoma; melanoma;
 myeloma; sarcoma.

OS Homo sapiens.
 XX
 XX WO9931128-A2.

24-JUN-1999.

02-DEC-1998; 98WO-US25649.

PR 16-DEC-1997; 97US-0991945.

XX (INCY-) INCYTE PHARM INC.

PI Au-Young J, Bandman O, Hillman JL, Kaser MR, Tang YT;

XX WPI; 1999-457916/38.

DR N-PSDB; X89503.

XX New tumour necrosis factor-R2-like protein - useful in the treatment
 PT of osteogenesis, developmental, reproductive, immunological and
 PT neoplastic disorders

PS Claim 1; Fig 1A-C; 8lpp; English.

XX The present sequence represents a human tumour necrosis factor-R2-like
 CC protein (TR2P). The protein is used to treat and prevent osteogenesis,
 CC developmental, reproductive, immunological and neoplastic disorders, and
 CC also to diagnose disorders associated with TR2 protein expression. Such
 CC disorders include osteogenesis disorders such as achondroplasia and
 CC osteoporosis, developmental disorders such as Cushing's syndrome,
 CC muscular dystrophy, and epilepsy, hereditary neuropathies such as
 CC Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism,
 CC hydrocephalus, seizure disorders such as cerebral palsy and spinal
 CC bifida, congenital glaucoma, cataract, or sensorineural hearing loss,
 CC reproductive disorders such as infertility, ovulatory defects and
 CC endometriosis, autoimmune disorders, ectopic pregnancy and teratogenesis,
 CC disruption of spermatogenesis, immunological disorders such as AIDS,
 CC Addison's disease, allergies, bronchitis, atherosclerosis, diabetes
 CC mellitus, Chron's disease, lupus and irritable bowel syndrome, multiple
 CC sclerosis, viral, fungal, helminthic, parasitic and protozoal infections,
 CC and neoplastic disorders including adenocarcinoma, leukaemia, lymphoma,
 CC melanoma, myeloma, sarcoma, and teratocarcinoma.

SQ Sequence 245 AA;

Query Match 100.0%; Score 59; DB 20; Length 245;
 Best Local Similarity 100.0%; Pred. No. 0.058;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HASCPPGAGV 10
 Db 129 hascppgagv 138
 |||||

RESULT 5
 Y42184
 ID Y42184 standard; Protein; 271 AA.

XX Y42184;

17-DEC-1999 (first entry)

Human mFLINT #1 protein sequence.

Human; FLINT; mFLINT; OP3; tumour necrosis factor receptor; FasL;
 apoptosis; inflammation; cancer; diabetes; acute liver failure;
 sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
 reperfusion-associated injury; aplastic anaemia; differentiation;
 growth; myelodysplastic syndrome; pancytopenic condition;
 myocardial ischaemia.

XX Homo sapiens.

XX WO9950413-A2.

07-OCT-1999.

30-MAR-1999; 99WO-US06797.

30-MAR-1998; 98US-0079856.

20-MAY-1998; 98US-0086074.

PN EP861850-A1.
 XX
 PD 02-SEP-1998.
 XX
 PF 20-JAN-1998; 98EP-0300382.
 XX
 PR 04-FEB-1997; 97US-0794796.
 XX
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 PA
 PI Emery J, Tan KB, Truneh A, Young PR;
 XX
 XX WPI; 1998-508248/44.
 DR
 DR N-PSDB; V07654.
 XX
 XX New DNA encoding tumour necrosis related receptor - used to treat
 PT and prevent e.g. inflammation, arthritis, septicemia, autoimmune
 PT diseases, transplant rejection, infection, stroke, ischaemia, ARDS,
 PT restenosis, AIDS, bone disorders and cancer
 XX
 PS Claim 1; Fig 1; 21pp; English.
 XX
 CC This is the amino acid sequence of the human tumour necrosis related
 CC receptor (TR4), used in the method of the invention. The TR4 protein
 CC or its agonist can be used to treat a subject in need of enhanced
 CC TR4 polypeptide activity. The antagonist is used to inhibit TR4
 CC polypeptide activity. The active agents can be used for the
 CC treatment and prevention of diseases such as chronic and acute
 CC inflammation, arthritis, septicemia, autoimmune diseases, transplant
 CC rejection, stroke, cancer, Alzheimer's disease.
 XX
 SQ Sequence 300 AA;
 Query Match 100.0%; Score 59; DB 19; Length 300;
 Best Local Similarity 100.0%; Pred. No. 0.07;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HASCPPGAGV 10
 Db 129 hascppgagv 138
 RESULT 8
 W63622
 ID W63622 standard; Protein; 300 AA.
 AC W63622;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Human tumour necrosis factor receptor-6 alpha protein.
 XX
 KW Human tumour necrosis factor receptor-6 alpha; TNFR-6 alpha; TNFR-6 beta;
 KW endothelial cells; keratinocytes; normal prostate; apoptosis;
 KW prostate tumour tissue.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..30
 FT Protein 31..300
 FT /note= "TNFR-6 alpha"
 FT Region 31..282
 FT /note= "Soluble extracellular domain"
 XX
 XX W09830694-A2.
 PN
 XX
 PD 16-JUL-1998.
 XX
 PF 13-JAN-1998; 98WO-US00153.
 XX
 XX 14-JAN-1997; 97US-0035496.
 PR

XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Ebner R, Feng P, Gentz RL, Ni J, Ruben SM, Yu G;
 PI
 XX WPI; 1998-399142/34.
 DR
 DR N-PSDB; V39085.
 XX
 XX Human tumour necrosis factor receptors 6-alpha and 6-beta - used in
 PT the diagnosis of immune system-related disorder(s)
 PT
 XX Claim 20; Fig 1; 91pp; English.
 PS
 XX
 CC The present sequence represents the human tumour necrosis factor
 CC receptor-6 alpha (TNFR-6 alpha) protein. The invention also provides
 CC for the TNFR-6 beta protein (W63623). TNFR-6 alpha and TNFR-6 beta are
 CC members of the tumour necrosis factor receptor (TNFR) family. TNFRs
 CC are expressed in endothelial cells, keratinocytes, normal prostate and
 CC prostate tumour tissue. For a number of disorders of these cells,
 CC particularly of the immune system, substantially altered (whether
 CC increased or decreased) levels of TNFR-6 alpha and/or TNFR-6 beta gene
 CC expression can be detected, therefore the TNFR-6 alpha and TNFR-6 beta
 CC polypeptides, nucleic acids and antibodies are claimed to be useful in
 CC the diagnosis of such disorders. Mutations of the TNFR-6 alpha and
 CC TNFR-6 beta genes can also be detected. The TNFR polypeptides are
 CC also claimed to be useful for identifying ligands which may be useful
 CC in the treatment of apoptosis related disorders.
 XX
 SQ Sequence 300 AA;
 Query Match 100.0%; Score 59; DB 19; Length 300;
 Best Local Similarity 100.0%; Pred. No. 0.07; Indels 0; Gaps 0;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HASCPPGAGV 10
 Db 129 hascppgagv 138
 RESULT 9
 Y03099
 ID Y03099 standard; Protein; 300 AA.
 XX
 AC Y03099;
 XX
 DT 09-DEC-1999 (first entry)
 DT
 DE Human lung TNF-receptor protein.
 DE
 XX
 KW Tumour necrosis factor; TNF; TNF receptor; human; lung; gene therapy;
 KW detection; immunoassay; diagnosis; disease; immune system; tumour;
 KW osteogenic system; cardiovascular system; central nervous system; asthma;
 KW peripheral nervous systems; transplant incompatibility; antitumor;
 KW rheumatoid arthritis; antiasthmatic; antiarthritic.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 134..1036
 FT /*tag= a
 FT /product= "TNF-receptor"
 XX
 XX DE19809978-A1.
 PN
 XX
 XX 16-SEP-1999.
 PD
 XX
 XX 09-MAR-1998; 98DE-1009978.
 PF
 XX
 XX 09-MAR-1998; 98DE-1009978.
 PR
 XX
 XX (BADI) BASF AG.
 PA
 XX

PI Kroeger B;
 DR WPI: 1999-519473/44.
 DR N-PSDB; Z09998.
 XX
 XX New soluble member of tumor necrosis factor receptor family, useful for
 PT identification specific modulators and for treating disease e.g. tumors
 PT
 XX
 PS Claim 1; Page 8-9; 10pp; German.
 XX
 CC This invention describes a novel tumour necrosis factor (TNF) receptor
 CC (I) isolated from human lung tissue. (I) is used: (i) to raise specific
 CC antibodies (Ab); (ii) to screen for specific (ant)agonists or ligands
 CC (A), potential therapeutic agents; and (iii) therapeutically (optionally
 CC expressed from a gene therapy vector) in conditions associated with a
 CC deficit of (I). Ab are used: (a) for qualitative or quantitative
 CC detection of (I) in standard immunoassays (for diagnosis of disease, or
 CC susceptibility, or for monitoring); and (b) as therapeutic inhibitors in
 CC cases where (I) is overexpressed. Nucleic acid (II) that encodes (I) is
 CC used: (A) for recombinant production of (I); (B) also its oligonucleotide
 CC fragments, in standard hybridization and/or amplification assays; (C) as
 CC source of antisense molecules or ribozymes; and (D) to produce transgenic
 CC animals (for studying (patho)physiology of (I)). Diseases possibly
 CC associated with under- or over-expression of (I) are those of the immune,
 CC osteogenic, cardiovascular and central or peripheral nervous systems,
 CC tumors, transplant incompatibility, asthma and rheumatoid arthritis. The
 CC products of the invention have antitumor, antiasthmatic and
 CC antiarthritic activity. This sequence represents the TNF-receptor of the
 CC invention.
 XX
 XX Sequence 300 AA;
 SQ
 Query Match 100.0%; Score 59; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 0.07;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HASCPPGAGV 10
 Db 129 hascppgagv 138
 |||||
 RESULT 10
 ID Y42182 standard; Protein; 300 AA.
 XX
 AC Y42182;
 XX
 XX 17-DEC-1999 (first entry)
 XX Human FLINT #1 protein sequence.
 XX Human; FLINT; mFLINT; OPC3; tumour necrosis factor receptor; FasL;
 KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
 KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
 KW reperfusion-associated injury; aplastic anaemia; differentiation;
 KW growth; myelodysplastic syndrome; pancytopenic condition;
 KW myocardial ischaemia.
 XX Homo sapiens.
 OS
 XX WO9950413-A2.
 PN
 XX 07-OCT-1999.
 PD
 XX 30-MAR-1999; 99WO-US06797.
 PF
 XX 30-MAR-1998; 98US-0079856.
 PR 20-MAY-1998; 98US-0086074.
 PR 09-SEP-1998; 98US-0099643.
 PR 17-DEC-1998; 98US-0112577.
 PR 18-DEC-1998; 98US-0112703.

PR 18-DEC-1998; 98US-0112933.
 PR 22-DEC-1998; 98US-0113407.
 XX
 PA (ELIL) LILLY & CO ELI.
 XX
 PI Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
 PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
 PI Song HY, Wang J, Wu X, Zuckerman SH;
 XX
 XX WPI: 1999-591319/50.
 DR N-PSDB; Z25375.
 DR
 XX Use of mature FLINT for treating acute liver failure, inflammation,
 PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
 PT and proinflammatory activity
 PT
 XX Claim 30; Fig 1; 90pp; English.
 PS
 XX The present invention describes therapeutic applications of mature FLINT
 CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
 CC (mFLINT), which is a member of the tumour necrosis factor receptor
 CC superfamily, is used for treating acute liver failure, inflammation of
 CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
 CC with inflammation, hepatitis, abnormal apoptosis (including use with
 CC injury or disorder such as hypercoagulation (including use with
 CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
 CC or disorder, Type I diabetes, cancer, cell damage or damage to an
 CC innocent bystander tissue that is induced by a chemotherapeutic agent or
 CC therapeutic irradiation, treating haematopoietic progenitor cells that
 CC have been exposed to therapeutic radiation or chemotherapy, aplastic
 CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
 CC also used for promoting the growth or differentiation of a haematopoietic
 CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
 CC resulting from abnormal myocardial ischaemia. The present sequence
 CC represents human FLINT.
 XX
 XX Sequence 300 AA;
 SQ
 Query Match 100.0%; Score 59; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 0.07;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HASCPPGAGV 10
 Db 129 hascppgagv 138
 |||||
 RESULT 11
 Y17479
 ID Y17479 standard; Protein; 300 AA.
 XX
 AC Y17479;
 XX
 XX 02-AUG-1999 (first entry)
 DT
 XX Mammalian tumour necrosis factor receptor OPC-2.
 DE
 XX Tumour necrosis factor receptor; TNF receptor; OPC-2; Paget's disease;
 KW osteopenic disorder; osteoclast activity; primary osteoporosis;
 KW hyperglycaemia; osteolytic metastasis; immune response; cancer.
 KW
 XX Mammalia.
 OS
 XX WO9926977-A1.
 PN
 XX 03-JUN-1999.
 PD
 XX 24-NOV-1998; 98WO-US25065.
 PF
 XX 17-FEB-1998; 98US-0074896.
 PR 24-NOV-1997; 97US-0066446.
 PR

PA (BIOJ) BIOGEN INC.
 XX
 PI Tschopp J;
 XX
 DR WPI: 1999-347693/29.
 DR N-PSDB: X76052.
 XX
 PT New tumour necrosis factor family receptor OPG-2
 XX
 XX
 PS Claim 1; Page 18; 22pp; English.
 XX
 CC The present sequence represents a mammalian tumour necrosis factor
 CC receptor, designated OPG-2. OPG-2, is a member of the tumour necrosis
 CC factor receptor family, and can be used: (i) to raise specific
 CC antibodies (Ab). (ii) to treat osteopenic disorders associated with
 CC excessive osteoclast activity, e.g. primary osteoporosis, Paget's
 CC disease, hyperglycemia of malignancy, or osteolytic metastases; (iii)
 CC for affinity purification of cognate ligands, and (iv) to screen for
 CC ligands (antagonists or agonists). Ab, or other OPG-2 blocking agents
 CC such as soluble forms of the protein, are used to prevent, or reduce
 CC severity of, an immune response, and for treating cancer. They can also
 CC be used in diagnostic assays. The nucleic acid sequence encoding OPG-2
 CC can be used as a probe to isolate related sequences from other species.
 XX
 SQ Sequence 300 AA;
 Query Match 100.0%; Score 59; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 0.07;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 HASCPPGAGV 10
 Db 129 hascppgagv 138
 RESULT 12
 Y06817
 ID Y06817 standard; Protein; 300 AA.
 XX
 AC Y06817;
 XX
 DT 24-JUN-1999 (first entry)
 XX
 DE Human Dcr3 polypeptide.
 XX
 KW Dcr3 polypeptide; tumour necrosis factor receptor; TNFR; Fas ligand;
 KW apoptosis; T cell mediated immune response; allergy; asthma; cancer;
 KW rheumatoid arthritis; Crohn's disease; guest vs. host disease; human;
 KW gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO9914330-A1.
 XX
 PD 25-MAR-1999.
 XX
 PF 18-SEP-1998; 98WO-US19661.
 XX
 PR 30-JUL-1998; 98US-0094640.
 PR 18-SEP-1997; 97US-0059288.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Ashkenazi AJ, Botstein D, Dodge KH, Goddard A, Gurney AL;
 PI Kim KJ, Lawrence DA, Pitti R, Roy MA, Tumas DB;
 PI Wood WI;
 XX
 DR WPI: 1999-244032/20.
 DR N-PSDB: X32744.
 XX
 PT Dcr3 polypeptide related to tumor necrosis factor receptor
 XX

PS Claim 5; Fig 1; 88pp; English.
 XX
 CC This represents a human Dcr3 polypeptide, a homologue of tumour necrosis
 CC factor receptor (TNFR) polypeptide. Host cells containing a vector
 CC comprising the Dcr3 nucleic acid can be used for the recombinant
 CC expression of the protein. Dcr3 binds to Fas ligand, so it (or its
 CC chimeras) are useful for modulating apoptosis in mammalian cells, also
 CC other Fas-ligand induced activities, particularly to inhibit T cell
 CC mediated immune responses, e.g. in treatment of allergy, asthma,
 CC rheumatoid arthritis, Crohn's disease, guest vs. host disease etc. Dcr3
 CC may also be used to identify specific binding proteins, potential
 CC inhibitors. Antibodies against Dcr3 are used to treat cancer,
 CC specifically of the lung and colon, also in diagnosis and for affinity
 CC purification of the protein. Detecting mutations in the gene for Dcr3 is
 CC also used to diagnose cancer, or predisposition to it. Dcr3 nucleic acid
 CC is useful as hybridization probe to detect genomic or related sequences;
 CC for chromosome and gene mapping; as source of antisense sequences;
 CC expression of recombinant Dcr3 and to generate transgenic animals (for
 CC development and screening of therapeutic agents), also for in vivo or
 CC ex vivo gene therapy.
 XX
 SQ Sequence 300 AA;
 Query Match 100.0%; Score 59; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 0.07;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 HASCPPGAGV 10
 Db 129 hascppgagv 138
 RESULT 13
 W97749
 ID W97749 standard; Protein; 300 AA.
 XX
 AC W97749;
 XX
 DT 21-MAY-1999 (first entry)
 XX
 DE Human tumour necrosis factor receptor 2TNFR-5.
 XX
 KW ZTNFR-5; tumour necrosis factor receptor; TNFR; human;
 KW cell maturation; bone cell regulation.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..23
 FT /note= "signal peptide"
 FT Protein 24..300
 FT /note= "mature protein"
 FT Domain 24..194
 FT /note= "extracellular domain"
 FT Region 49..71
 FT /note= "cysteine-rich pseudo-repeat 1"
 FT Region 72..113
 FT /note= "cysteine-rich pseudo-repeat 1"
 FT Region 114..151
 FT /note= "cysteine-rich pseudo-repeat 1"
 FT Region 152..194
 FT /note= "cysteine-rich pseudo-repeat 1"
 XX
 PN WO9904001-A1.
 XX
 PD 28-JAN-1999.
 XX
 PF 21-JUL-1998; 98WO-US15072.
 XX
 PR 21-JUL-1997; 97US-0053203.
 XX
 PA (ZYMO) ZYMOGENETICS INC.

```

XX PI Farrah TM;
XX DR WPI; 1999-132245/11.
XX DR N-PSDB; X07226.
XX PT Novel tumour necrosis factor receptor ZTNFR5 - useful for
XX PT regulating maturation of TNF-ligand bearing cells
XX PS Claim 1; Page 84-85; 109pp; English.
XX CC This polypeptide comprises a new, secreted tumour necrosis factor
XX CC receptor (see W97749), designated ZTNFR-5. Novel ZTNFR-5 encoding
XX CC polynucleotides and polypeptides were initially identified by
XX CC querying an expressed sequence tag (EST) database for sequences
XX CC homologous to conserved motifs within the TNF receptor family.
XX CC Based on this search, a contig of 16 ESTs (see X07226) was
XX CC constructed. ZTNFR-5 polypeptides comprise 4 cysteine-rich repeats
XX CC (see also W97750-55) that are homologous to other TNF receptors, in
XX CC particular the soluble, secreted TNF receptor osteoprotegerin.
XX CC ZTNFR-5 polypeptide can be prepared by recombinant methods. The
XX CC polypeptide, especially the extracellular domain, can be used to
XX CC generate a soluble variant of ZTNFR-5. The polypeptides and
XX CC nucleic acids can be used to screen for ligands, agonists and
XX CC antagonists of ZTNFR-5. The polypeptides can be used in bone cell
XX CC regulation and to regulate the maturation of TNF ligand-bearing
XX CC cells such as T- or B-cells, lymphocytes, peripheral blood
XX CC mononuclear cells, polymorphonuclear leukocytes, fibroblasts or
XX CC haematopoietic cells.
XX SQ Sequence 300 AA;

Query Match 100.0%; Score 59; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 0.07;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HASCPPGAGV 10
Db 129 hascppgagv 138
|||||

RESULT 14
W95082
ID W95082 standard; Protein; 300 AA.
XX AC W95082;
XX DT 20-MAY-1999 (first entry)
XX DE Orphan receptor (HUMAN NTR-1) polypeptide.
XX KW HUMAN NTR-1; orphan receptor; osteoprotegerin; OPG; TNFR; human;
XX KW tumour necrosis factor receptor; muscle disorder; bone mass; screening;
XX KW muscle metabolism; binding agent; cognate ligand.
XX OS Homo sapiens.
XX PN W09907738-A2.
XX PD 18-FEB-1999.
XX PF 04-AUG-1998; 98WO-US16202.
XX PR 06-AUG-1997; 97US-0054869.
XX XX (PROC ) PROCTER & GAMBLE CO.
XX PA (REG-) REGENERON PHARM INC.
XX PI Maslowski PJ, Morris J, Valenzuela DM;
XX DR WPI; 1999-167365/14.
XX DR N-PSDB; X22300.

Query Match 100.0%; Score 59; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 0.07;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HASCPPGAGV 10
Db 129 hascppgagv 138
|||||

RESULT 15
Y77458
ID Y77458 standard; Protein; 300 AA.
XX AC Y77458;
XX DT 05-JUN-2000 (first entry)
XX DE Human TNF receptor-like protein, HDTEA84.
XX KW TNF receptor family; tumour necrosis factor; HDTEA84; HSLJD37R;
XX KW Rank-like protein; RANKL; immune disorder; inflammation; allergy;
XX KW immunosuppressant; antiarthritic; antirheumatoid; antiinflammatory;
XX KW dermatological; antithyroid.
XX OS Homo sapiens.
XX PN W0200001817-A2.
XX PD 13-JAN-2000.
XX PF 06-JUL-1999; 99WO-US12366.
XX PR 06-JUL-1998; 98US-0110938.
XX PR 13-JUL-1998; 98US-0114466.
XX PR 23-JUL-1998; 98US-0093897.
XX PR 12-AUG-1998; 98US-0132968.
XX PR 18-AUG-1998; 98US-0136214.
XX PR 11-SEP-1998; 98US-0099999.
XX PA (SCHE ) SCHERING CORP.
XX PI Bates EEM, Lebecque SJE, Murphy EE, Mattson JD, Gorman DM;
XX PI Hedrick JA, Wang L, Zlotnik A, Murgolo NJ, Greene JR, Johnston JA;
XX PI Bazan JF, Mahony D, Lees EM;
XX DR WPI; 2000-171015/15.
XX DR N-PSDB; Z92404.
XX XX New isolated mammalian genes, used to develop products for treating
XX PT e.g. immune, inflammatory or allergic abnormalities, cancers or
XX PT degenerative conditions
XX PS Claim 24; Page 157; 218pp; English.

```

```

XX PI Farrah TM;
XX DR WPI; 1999-132245/11.
XX DR N-PSDB; X07226.
XX PT Novel tumour necrosis factor receptor ZTNFR5 - useful for
XX PT regulating maturation of TNF-ligand bearing cells
XX PS Claim 1; Page 84-85; 109pp; English.
XX CC This polypeptide comprises a new, secreted tumour necrosis factor
XX CC receptor (see W97749), designated ZTNFR-5. Novel ZTNFR-5 encoding
XX CC polynucleotides and polypeptides were initially identified by
XX CC querying an expressed sequence tag (EST) database for sequences
XX CC homologous to conserved motifs within the TNF receptor family.
XX CC Based on this search, a contig of 16 ESTs (see X07226) was
XX CC constructed. ZTNFR-5 polypeptides comprise 4 cysteine-rich repeats
XX CC (see also W97750-55) that are homologous to other TNF receptors, in
XX CC particular the soluble, secreted TNF receptor osteoprotegerin.
XX CC ZTNFR-5 polypeptide can be prepared by recombinant methods. The
XX CC polypeptide, especially the extracellular domain, can be used to
XX CC generate a soluble variant of ZTNFR-5. The polypeptides and
XX CC nucleic acids can be used to screen for ligands, agonists and
XX CC antagonists of ZTNFR-5. The polypeptides can be used in bone cell
XX CC regulation and to regulate the maturation of TNF ligand-bearing
XX CC cells such as T- or B-cells, lymphocytes, peripheral blood
XX CC mononuclear cells, polymorphonuclear leukocytes, fibroblasts or
XX CC haematopoietic cells.
XX SQ Sequence 300 AA;

Query Match 100.0%; Score 59; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 0.07;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HASCPPGAGV 10
Db 129 hascppgagv 138
|||||

RESULT 14
W95082
ID W95082 standard; Protein; 300 AA.
XX AC W95082;
XX DT 20-MAY-1999 (first entry)
XX DE Orphan receptor (HUMAN NTR-1) polypeptide.
XX KW HUMAN NTR-1; orphan receptor; osteoprotegerin; OPG; TNFR; human;
XX KW tumour necrosis factor receptor; muscle disorder; bone mass; screening;
XX KW muscle metabolism; binding agent; cognate ligand.
XX OS Homo sapiens.
XX PN W09907738-A2.
XX PD 18-FEB-1999.
XX PF 04-AUG-1998; 98WO-US16202.
XX PR 06-AUG-1997; 97US-0054869.
XX XX (PROC ) PROCTER & GAMBLE CO.
XX PA (REG-) REGENERON PHARM INC.
XX PI Maslowski PJ, Morris J, Valenzuela DM;
XX DR WPI; 1999-167365/14.
XX DR N-PSDB; X22300.

Query Match 100.0%; Score 59; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 0.07;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HASCPPGAGV 10
Db 129 hascppgagv 138
|||||

RESULT 15
Y77458
ID Y77458 standard; Protein; 300 AA.
XX AC Y77458;
XX DT 05-JUN-2000 (first entry)
XX DE Human TNF receptor-like protein, HDTEA84.
XX KW TNF receptor family; tumour necrosis factor; HDTEA84; HSLJD37R;
XX KW Rank-like protein; RANKL; immune disorder; inflammation; allergy;
XX KW immunosuppressant; antiarthritic; antirheumatoid; antiinflammatory;
XX KW dermatological; antithyroid.
XX OS Homo sapiens.
XX PN W0200001817-A2.
XX PD 13-JAN-2000.
XX PF 06-JUL-1999; 99WO-US12366.
XX PR 06-JUL-1998; 98US-0110938.
XX PR 13-JUL-1998; 98US-0114466.
XX PR 23-JUL-1998; 98US-0093897.
XX PR 12-AUG-1998; 98US-0132968.
XX PR 18-AUG-1998; 98US-0136214.
XX PR 11-SEP-1998; 98US-0099999.
XX PA (SCHE ) SCHERING CORP.
XX PI Bates EEM, Lebecque SJE, Murphy EE, Mattson JD, Gorman DM;
XX PI Hedrick JA, Wang L, Zlotnik A, Murgolo NJ, Greene JR, Johnston JA;
XX PI Bazan JF, Mahony D, Lees EM;
XX DR WPI; 2000-171015/15.
XX DR N-PSDB; Z92404.
XX XX New isolated mammalian genes, used to develop products for treating
XX PT e.g. immune, inflammatory or allergic abnormalities, cancers or
XX PT degenerative conditions
XX PS Claim 24; Page 157; 218pp; English.

```


xx The invention relates to a number of primate and/or rodent proteins, and
 CC the genes which encode them. The invention encompasses human dendritic
 CC cell prostaglandin transporter (DC-PGT); the TNF (tumour necrosis
 CC factor) receptor family-related proteins HDTEA84, HSLJD37R and RANKL;
 CC human CC chemokine HCC5; human deubiquitinating proteins Dub11 and Dub
 CC 12; human MD-1 and human and murine MD-2 proteins, which exhibit the
 CC properties of ligands for proteins comprising a leucine-rich motif (LRR);
 CC human cyclin E2; cDNAs encoding these proteins; and antibodies against
 CC these proteins. The proteins can be used for modulating the physiology or
 CC development of a cell. They can be used to mediate uptake of substrates
 CC (e.g., prostaglandin-like molecules), to modulate or mediate cellular
 CC interactions (e.g., induce or prevent trafficking, proliferation, or
 CC differentiation of cells), or are intracellular proteins which are
 CC important in various cellular processes such as the deubiquitination of
 CC proteins or cell cycle regulation. The products can be used for
 CC treating medical conditions such as immune, inflammatory or allergic
 CC disorders, or abnormal cellular proliferation, for example, cancers or
 CC degenerative conditions. They can be used to modulate immune responses in
 CC disease states e.g., autoimmune disorders, including rheumatoid
 CC arthritis, systemic lupus erythematosus, Hashimoto's autoimmune
 CC thyroiditis, as well as acute and chronic inflammatory responses in which
 CC T cell activation, expansion, and/or immunological T cell memory play an
 CC important role. Sequences Y77458-Y77461 and Y77465-Y77468 represent
 CC TNF receptor family-related proteins. Y77458 is the human protein
 CC HDTEA84, Y77459-Y77461 are human HSLJD37R proteins, Y77465 is
 CC murine Rank-like protein RANKL, and Y77466-Y77468 are human RANKL
 CC proteins.

xx
 SQ Sequence 300 AA;

Query Match 100.0%; Score 59; DB 21; Length 300;
 Best Local Similarity 100.0%; Pred. No. 0.07; 0; Indels 0; Gaps 0;
 Matches 10; Conservative 0; Mismatches 0;

Oy 1 HASCPPGAGV 10
 |||||
 Db 129 hascppgagv 138

Search completed: January 30, 2001, 16:45:31
 Job time: 591 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:50:38 ; Search time 149.64 seconds
(without alignments)
4,538 Million cell updates/sec

Title: US-09-518-931-4_COPY_129_138

Perfect score: 59

Sequence: 1 HASCPPGAGV 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_66: *
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	69.5	299	2 A75591	hypothetical prote
2	40	67.8	277	2 A60771	B-cell activation
3	40	67.8	278	2 C69506	probable 2-oxoisov
4	40	67.8	780	2 A34102	von Willebrand fac
5	40	67.8	2813	1 VWHU	hypothetical prote
6	39	66.1	395	2 G75587	fibulin-2 precursor
7	39	66.1	1184	2 A55184	fibulin-2 precursor
8	39	66.1	1221	2 A49457	ATP binding casset
9	39	66.1	2201	2 A54774	probable cobM - My
10	38	64.4	251	2 B70765	3-deoxy-manno-octu
11	38	64.4	427	2 F64084	hypothetical prote
12	38	64.4	480	2 H70854	M-sema F protein p
13	38	64.4	834	2 S66498	progesterone recep
14	38	64.4	933	1 QRHP	hypothetical prote
15	38	64.4	981	2 T46330	notch4 - mouse
16	38	64.4	1964	2 T09059	notch3 protein - h
17	38	64.4	2321	2 S78549	cation-independent
18	38	64.4	2470	2 I50726	laminin alpha 5 ch
19	38	64.4	3635	2 T10053	hypothetical prote
20	37	62.7	73	2 JS0514	hypothetical prote
21	37	62.7	248	2 T2632	precorrin-3 methyl
22	37	62.7	261	2 T03537	hypothetical prote
23	37	62.7	268	2 S77310	regulatory protein
24	37	62.7	332	2 I48691	gag polypeptide -
25	37	62.7	392	1 FOLJCA	gag protein - bovi
26	37	62.7	393	2 S29356	hypothetical prote
27	37	62.7	413	2 A70795	hypothetical prote
28	37	62.7	473	2 C70915	protein-tyrosine k
29	37	62.7	517	2 S24547	

30	37	62.7	548	1 UFECAQ	fumarate hydratase
31	37	62.7	548	1 B44511	fumarate hydratase
32	37	62.7	572	2 T47219	amino acid transpo
33	37	62.7	585	2 A39395	delayed rectifier
34	37	62.7	650	2 T00094	endostyle-specific
35	37	62.7	994	2 T49276	c-mer tyrosine kin
36	37	62.7	1024	2 T16491	hypothetical prote
37	37	62.7	1101	2 T16840	hypothetical prote
38	37	62.7	2224	1 KPHU5	coagulation factor
39	36	61.0	22	2 S64675	collagen alpha 3(I
40	36	61.0	339	2 S09284	Ig alpha chain C r
41	36	61.0	565	2 T16408	hypothetical prote
42	36	61.0	880	2 T02245	hypothetical prote
43	36	61.0	2195	2 T34264	hypothetical prote
44	36	61.0	2318	2 S45306	notch 3 protein -
45	36	61.0	2491	1 A28372	insulin-like growt

ALIGNMENTS

RESULT 1

A75591

hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C:Accession: A75591

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J
S.; Smith, H.O.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896

A:Accession: A75591

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-299 <WHI>

A:Cross-references: GB:A6001863; GB:A6001825; NID:g6460670; PIDN:AAF12505.1; PID:g646

A:Experimental source: strain R1

C:Genetics:

A:Gene: DRA0358

A:Map position: 2

Query Match 69.5%; Score 41; DB 2; Length 299;

Best Local Similarity 72.7%; Pred. No. 21;

Matches 8; Conservative 1; Mismatches 0; Indels 2; Gaps 1;

Oy 1 HASC--PPGAG 9

Db 97 HAACLTLPFGAG 107

RESULT 2

A60771

B-cell activation protein CD40 precursor - human

N:Alternate names: B-cell surface antigen Bp50

C:Species: Homo sapiens (man)

C:Date: 03-Jun-1993 #sequence_revision 03-Feb-1994 #text_change 21-Jul-2000

C:Accession: S04460; A60771

R:Stamenkovic, I.; Clark, E.A.; Seed, B.

EMBO J. 8, 1403-1410, 1989

A:Title: A B-lymphocyte activation molecule related to the nerve growth factor recept

A:Reference number: S04460; MUID:89356608

A:Accession: S04460

A:Molecule type: mRNA

A:Residues: 1-277 <STA>

A:Cross-references: EMBL:X60592; NID:g29850; PIDN:CAA43045.1; PID:g29851

R:Braesch-Andersen, S.; Paulie, S.; Koho, H.; Aspenstrom, P.; Perlmann, P.
J. Immunol. 142, 562-567, 1989

A:Title: Biochemical characteristics and partial amino acid sequence of the receptor-

A:Reference number: A60771; MUID:89093941

A:Accession: A60771

A:Molecule type: protein
A:Residues: 21-50

A:Experimental source: Burkitt lymphoma cell line Raji
C:Genetics:
A:Gene: GDB:CD40
A:Cross-references: GDB:215268; OMIM:109535
A:Map position: 20q12-20q13.2
C:Superfamily: CD27 antigen; NGF receptor repeat homology
C:Keywords: B-cell; glycoprotein; phosphoprotein; surface antigen; transmembrane protein
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-277/Product: B-cell activation protein CD40 #status experimental <MAT>
F:21-193/Domain: extracellular #status predicted <EXT>
F:194-215/Domain: transmembrane #status predicted <TM>
F:216-277/Domain: intracellular #status predicted <CYT>
F:153,180/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 67.8%; Score 40; DB 2; Length 277;
Best Local Similarity 70.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 HASCPPGAGV 10
| | | | |
Db 122 HRSCPPGFGV 131

RESULT 3
C69506
probable 2-oxoisovalerate--ferredoxin oxidoreductase (EC 1.2.7.-) beta chain AF2052 [sim
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000
C:Accession: C69506
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiaich, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A:Reference number: A69250; MUID:98049343
A:Accession: C69506
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-278 <KLE>
A:Cross-references: GB:AE000961; GB:AE000782; NID:g2689284; PIDN:AAB89197.1; PID:g264847
C:Superfamily: pyruvate synthase beta chain
C:Keywords: oxidoreductase

Query Match 67.8%; Score 40; DB 2; Length 278;
Best Local Similarity 85.7%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HASCPPG 7
| | | | |
Db 189 HAPCPPG 195

RESULT 4
A34102
von Willebrand factor - bovine (fragments)
C:Species: Bos primigenius taurus (cattle)
C:Date: 30-Mar-1990 #sequence_revision 05-May-1995 #text_change 30-Jan-1998
C:Accession: A34102; S14768; A39540; S27196; S36600; S68346
R:Takagi, J.; Kasahara, K.; Sekiya, F.; Inada, Y.; Saito, Y.
J. Biol. Chem. 264, 10425-10430, 1989
A:Title: A collagen-binding glycoprotein from bovine platelets is identical to propolype
A:Reference number: A34102; MUID:89278101
A:Accession: A34102
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-21; 78-98 <PAK>
R:Fujisawa, T.; Takagi, J.; Sekiya, F.; Goto, A.; Miake, F.; Saito, Y.

Eur. J. Biochem. 196, 673-677, 1991
A:Title: Monoclonal antibodies that inhibit binding of propolypeptide of von Willebra
A:Reference number: S14768; MUID:91192039
A:Accession: S14768
A:Status: preliminary
A:Molecule type: protein
A:Residues: 22-51; 52-77; 112-149 <FUJ>
R:Takagi, J.; Fujisawa, T.; Sekiya, F.; Saito, Y.
J. Biol. Chem. 266, 5575-5579, 1991
A:Title: Collagen-binding domain within bovine propolypeptide of von Willebrand facto
A:Reference number: A39540; MUID:91170224
A:Accession: A39540
A:Status: preliminary
A:Molecule type: protein
A:Residues: 99-257; 258-264 <TA2>
R:Bakshi, M.R.; Myers, J.C.; Howard, P.S.; Soprano, D.R.; Kirby, E.P.
Biochim. Biophys. Acta 1132, 325-328, 1992
A:Title: Sequencing of the primary adhesion domain of bovine von Willebrand factor.
A:Reference number: S27196; MUID:93041938
A:Accession: S27196
A:Molecule type: mRNA
A:Residues: 265-780 <BAK>
A:Cross-references: EMBL:X63820
R:Bakshi, M.R.
submitted to the EMBL Data Library, January 1992
A:Reference number: S36600
A:Accession: S36600
A:Molecule type: mRNA
A:Residues: 265-615; 'SRVA' 620-687, 'M' 689-718, 'G' 720-780 <BA2>
A:Cross-references: EMBL:X63820
R:Takagi, J.; Aoyama, T.; Ueki, S.; Ohba, H.; Saito, Y.; Lorand, L.
Eur. J. Biochem. 232, 773-777, 1995
A:Title: Identification of factor-XIIa-reactive glutamyl residues in the propolype
A:Reference number: S68346; MUID:96028118
A:Accession: S68346
A:Molecule type: protein
A:Residues: 22-33; LSPVYAGRTGCLGNV'; 99-101, 'X' 103-122 <TAI>
C:Superfamily: von Willebrand factor; von Willebrand factor type A repeat homology; v
C:Keywords: blood coagulation; cell binding; connective tissue; extracellular matrix;
F:513-681/Domain: von Willebrand factor type A repeat homology <WVAI>
F:755/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 67.8%; Score 40; DB 2; Length 780;
Best Local Similarity 85.7%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HASCPPG 7
| | | | |
Db 414 HAHCPPG 420

RESULT 5
VWHU
von Willebrand factor precursor - human
C:Species: Homo sapiens (man)
C:Date: 04-Dec-1986 #sequence_revision 30-Jun-1993 #text_change 22-Jun-1999
C:Accession: A34480; S02377; A37139; S23676; A25298; A25366; S23618; S23645;
R:Mancuso, D.J.; Tuley, E.A.; Westfield, L.A.; Worrall, N.K.; Shelton-Inloes, B.B.; S
J. Biol. Chem. 264, 19514-19527, 1989
A:Title: Structure of the gene for human von Willebrand factor.
A:Reference number: A34480; MUID:90062044
A:Accession: A34480
A:Molecule type: DNA
A:Residues: 1-2813 <MAN>
A:Cross-references: EMBL:M25864
R:Bonthron, D.; Orkin, S.H.
Eur. J. Biochem. 171, 51-57, 1988
A:Title: The human von Willebrand factor gene. Structure of the 5' region.
A:Reference number: S02377; MUID:88111704
A:Accession: S02377
A:Molecule type: DNA
A:Residues: 1-177 <BO2>

A:Cross-references: EMBL:X06828
 R:Mancuso, D.J.; Tuley, E.A.; Westfield, L.A.; Lester-Mancuso, T.L.; Le Beau, M.M.; Sora Biochemistry 30, 253-269, 1991
 A:Title: Human von Willebrand factor gene and pseudogene: structural analysis and differed
 A:Reference number: A37139; MUID:91105089
 A:Accession: A37139
 A:Molecule type: DNA
 A:Residues: 990-1947 <MAD>
 A:Cross-references: GB:M60675; NID:g340357; PIDN:AAA61295.1; PID:g553810
 A:Note: the authors translated the codon CGC for residue 156 as Gln
 R:Collins, C.J.; Underdahl, J.P.; Levene, R.B.; Ravera, C.P.; Morin, M.J.; Dombalagian, Proc. Natl. Acad. Sci. U.S.A. 84, 4393-4397, 1987
 A:Title: Molecular cloning of the human gene for von Willebrand factor and identification
 A:Reference number: S23676; MUID:87260814
 A:Accession: S23676
 A:Molecule type: DNA
 A:Residues: 2731-2813 <COL>
 A:Cross-references: EMBL:M16945
 R:Bonthron, D.; Orr, E.C.; Mitsock, L.M.; Ginsburg, D.; Handin, R.I.; Orkin, S.H. Nucleic Acids Res. 14, 7125-7127, 1986
 A:Title: Nucleotide sequence of pre-pro-von Willebrand factor cDNA.
 A:Reference number: A25298; MUID:87016349
 A:Accession: A25298
 A:Molecule type: mRNA
 A:Residues: 1-470, 'V', 472-2813 <BON>
 R:Verweij, C.L.; Diergaarde, P.J.; Hart, M.; Pannekoek, H. EMBO J. 5, 1839-1847, 1986
 A:Title: Full-length von Willebrand factor (vWF) cDNA encodes a highly repetitive protein
 A:Reference number: A91044; MUID:87004550
 A:Accession: A25469
 A:Molecule type: mRNA
 A:Residues: 1-470, 'V', 472-483, 'R', 485-1022, 'K', 1024-1025, 'E', 1027-1400 <VER>
 A:Cross-references: EMBL:X04146
 R:Verweij, C.L.; Diergaarde, P.J.; Hart, M.; Pannekoek, H. EMBO J. 5, 3074, 1986
 A:Reference number: A91056
 A:Accession: A25366
 A:Molecule type: mRNA
 A:Residues: 1021-1030 <VE2>
 A:Note: this is a revision to the sequence from reference A91044
 R:Shelton-Inloes, B.B.; Broze Jr., G.J.; Miletich, J.P.; Sadler, J.E. Biochem. Biophys. Res. Commun. 144, 657-665, 1987
 A:Title: Evolution of human von Willebrand factor: cDNA sequence polymorphisms, repeated
 A:Reference number: S23618; MUID:87213253
 A:Accession: S23618
 A:Molecule type: mRNA
 A:Residues: 1-120 <SH2>
 A:Cross-references: EMBL:M17588; NID:g799330; PIDN:AAA65940.1; PID:g340316
 A:Accession: S23645
 A:Molecule type: protein
 A:Residues: 23-56 <SH3>
 R:Sadler, J.E.; Shelton-Inloes, B.B.; Sorace, J.M.; Harlan, J.M.; Titani, K.; Davie, E.W. Proc. Natl. Acad. Sci. U.S.A. 82, 6394-6398, 1985
 A:Title: Cloning and characterization of two cDNAs coding for human von Willebrand factor
 A:Reference number: A94060; MUID:86016708
 A:Accession: A94060
 A:Molecule type: mRNA
 A:Residues: 'WA', 739, 'C', 744-769, 'H', 771-788, 'A', 790-803, 'S', 805-873; 1289-1471, 'D', 1473-1485
 A:Note: the authors translated the codon TCG for residue 2168 as Cys
 R:Shelton-Inloes, B.B.; Titani, K.; Sadler, J.E. Biochemistry 25, 3164-3171, 1986
 A:Title: cDNA sequences for human von Willebrand factor reveal five types of repeated do
 A:Reference number: A90504; MUID:86269894
 A:Accession: A90504
 A:Molecule type: mRNA
 A:Residues: 781-788, 'A', 790-1424 <SHE>
 A:Note: 852-Gln, 857-Asp, and 1381-Thr were also found
 R:Ginsburg, D.; Handin, R.I.; Bonthron, D.T.; Donlon, T.A.; Bruns, G.A.P.; Latt, S.A.; C Science 228, 1401-1406, 1985
 A:Title: Human von Willebrand factor (vWF): isolation of complementary DNA (cDNA) clones
 A:Reference number: A44178; MUID:85244588

A:Accession: A44178
 A:Molecule type: mRNA
 A:Residues: 2621-2813 <GIN>
 A:Cross-references: EMBL:X03028; NID:g340308; PIDN:AAA61293.1; PID:g340309
 R:Verweij, C.L.; de Vries, C.J.M.; Distel, B.; van Zonneveld, A.J.; van Kessel, A.G.; Nucleic Acids Res. 13, 4699-4717, 1985
 A:Title: Construction of cDNA coding for human von Willebrand factor using antibody p
 A:Reference number: S07363; MUID:85269603
 A:Accession: S07363
 A:Molecule type: mRNA
 A:Residues: 2731-2813 <VE3>
 A:Cross-references: EMBL:X02672; NID:g37939; PIDN:CAA26503.1; PID:g37940
 R:Lynch, D.C.; Zimmermann, T.S.; Collins, C.J.; Brown, M.; Morin, M.J.; Ling, E.H.; Li Cell 41, 49-56, 1985
 A:Title: Molecular cloning of cDNA for human von Willebrand factor: authentication by
 A:Reference number: S23678; MUID:85201687
 A:Accession: S23678
 A:Molecule type: mRNA
 A:Residues: 2731-2813 <LYN>
 A:Cross-references: EMBL:X03028
 R:Titani, K.; Kumar, S.; Takio, K.; Ericsson, L.H.; Wade, R.D.; Ashida, K.; Walsh, K. Biochemistry 25, 3171-3184, 1986
 A:Title: Amino acid sequences of human von Willebrand factor.
 A:Reference number: A90505; MUID:86269895
 A:Accession: A90505
 A:Molecule type: protein
 A:Residues: 764-788, 'A', 790-1471, 'D', 1473-2813 <TIT>
 A:Note: 789-Thr was also found
 R:Chopek, M.W.; Gilma, J.P.; Fujikawa, K.; Davie, E.W.; Titani, K. Biochemistry 25, 3146-3155, 1986
 A:Title: Human von Willebrand factor: a multivalent protein composed of identical sub
 A:Reference number: A23464; MUID:86269892
 A:Accession: A23464
 A:Molecule type: protein
 A:Residues: 764-773; 2803-2813 <CHO>
 R:Dent, J.A.; Berkowitz, S.D.; Ware, J.; Kasper, C.K.; Ruggeri, Z.M. Proc. Natl. Acad. Sci. U.S.A. 87, 6306-6310, 1990
 A:Title: Identification of a cleavage site directing the immunochemical detection of
 A:Reference number: A36013; MUID:90349604
 A:Accession: A36013
 A:Molecule type: protein
 A:Residues: 1606-1617 <DEN>
 R:Fay, P.J.; Kawai, Y.; Wagner, D.D.; Ginsburg, D.; Bonthron, D.; Ohlsson-Wilhelm, B. Science 232, 995-998, 1986
 A:Title: Propolypeptide of von Willebrand factor circulates in blood and is identical
 A:Reference number: A60913; MUID:86208144
 A:Accession: A60913
 A:Molecule type: protein
 A:Residues: 576-590 <FAY>
 C:Genetics:
 A:Gene: GDB:VWF
 A:Cross-references: GDB:119125; OMIM:193400
 A:Map position: 12p13.3-12p13.2
 A:Introns: 19/1; 74/1; 108/2; 178/1; 219/3; 292/1; 333/1; 370/2; 386/1; 431/3; 478/1; 5/1; 1724/1; 1771/1; 1819/1; 1874/1; 1888/3; 1948/1; 2021/3; 2086/1; 2200/1; 2266/3;
 C:Superfamily: von Willebrand factor; von Willebrand factor type A repeat homology; v
 C:Keywords: blood coagulation; cell binding; connective tissue; disulfide bond; dupli
 F:1-22/Domain: signal sequence #status predicted <Sig>
 F:23-763/Product: von Willebrand antigen II #status predicted <MA1>
 F:334-386/Domain: type D repeat 1 <DD1>
 F:387-745/Domain: type D repeat 2 <DD2>
 F:698-700/Region: cell attachment (R-G-D) motif
 F:764-2813/Product: von Willebrand factor #status predicted <MA2>
 F:784-865/Domain: 'D' <DDD>
 F:788-833, 2216-2261/Region: duplication
 F:826-853, 2400-2515, 2544-2662/Region: duplication
 F:842-1130, 1934-2203/Region: duplication
 F:866-1241/Domain: type D repeat 3 <DD3>
 F:1275-1443/Domain: von Willebrand factor type A repeat homology <VWA1>
 F:1496-1654/Domain: von Willebrand factor type A repeat homology <VWA2>
 F:1689-1854/Domain: von Willebrand factor type A repeat homology <VWA3>
 F:1947-2295/Domain: type D repeat 4 <DD4>
 F:2296-2330/Domain: type B repeat 1 <VB1>

F:2340-2365/Domain: type B repeat 2 <VB2>
 F:2375-2399/Domain: type B repeat 3 <VB3>
 F:2430-2497/Domain: von Willebrand factor type C repeat homology <VWC1>
 F:2507-2509/Region: cell attachment (R-G-D) motif
 F:2581-2647/Domain: von Willebrand factor type C repeat homology <VWC2>
 F:857,1231,1515,1574,2223,2290,2357,2400,2546,2585,2790/Binding site: carbohydrate (Asn)
 F:1147/Binding site: carbohydrate (Asn) (covalent) #status atypical
 F:1248,1255,1256,1468,1477,1487,1679,2298/Binding site: carbohydrate (Thr) (covalent) #s
 F:1263,1486/Binding site: carbohydrate (Ser) (covalent) #status experimental

Query Match 67.8%; Score 40; DB 1; Length 2813;
 Best Local Similarity 85.7%; Pred. No. 2.1e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HASCPPG 7
 || ||||
 Db 1174 HAHCPPG 1180

RESULT 6
 G75587
 hypothetical protein - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C:Accession: G75587
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896
 A:Accession: G75587
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-395 <WHI>
 A:Cross-references: GB:AE001863; GB:AE001825; NID:g6460670; PID:AAFI2510.1; PID:g646080
 A:Experimental source: strain R1
 C:Genetics: DRA0330
 A:Gene: DRA0330
 A:Map position: 2

Query Match 66.1%; Score 39; DB 2; Length 395;
 Best Local Similarity 87.5%; Pred. No. 55;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ASCPPGAG 9
 |||||
 Db 49 ASCPPGAG 56

RESULT 7
 A55184
 fibulin-2 precursor - human
 N:Alternate names: protein DKFzp586A1519.1
 C:Species: Homo sapiens (man)
 C>Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 21-Jul-2000
 C:Accession: A55184; T08744
 R:Zhang, R.Z.; Pan, T.C.; Zhang, Z.Y.; Mattei, M.G.; Timpl, R.; Chu, M.L.
 Genomics 22, 425-430, 1994
 A:Title: Fibulin-2 (FBLN2): human cDNA sequence, mRNA expression, and mapping of the gene
 A:Reference number: A55184; MUID:95104855
 A:Accession: A55184
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1184 <ZHA>
 A:Cross-references: GB:X82494; NID:g575232; PID:CAA57876.1; PID:g575233
 R:Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, March 1999

A:Reference number: Z16471
 A:Accession: T08744
 A:Molecule type: mRNA

A:Residues: 656-719, 'QDECLMGADCSRRQFCVNTLGSFYCNHTVLCADGYILNAHRKCVD', 720-853, 'T', 855
 A:Cross-references: EMBL:AL050095
 A:Experimental source: adult uterus; clone DKFzp586A1519
 C:Genetics:

A:Gene: GDB:FBLN2
 A:Cross-references: GDB:293037; OMIM:135821
 A:Map position: 3p25-3p24
 A>Note: DKFzp586A1519.1
 C:Superfamily: unassigned EGF-related proteins; EGF homology
 C:Keywords: alternative splicing; extracellular matrix
 F:1-27/Domain: signal sequence #status predicted <SIG>
 F:28-1184/Product: fibulin-2 protein #status predicted <MAT>
 F:905-941/Domain: EGF homology <EGF>

Query Match 66.1%; Score 39; DB 2; Length 1184;
 Best Local Similarity 85.7%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 SCPPGAG 9
 |||||
 Db 99 SCPPGGG 105

RESULT 8
 A49457
 fibulin-2 precursor - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 11-Jan-2000
 C:Accession: A49457; S74095
 R:Pan, T.C.; Sasaki, T.; Zhang, R.Z.; Faessler, R.; Timpl, R.; Chu, M.L.
 J. Cell Biol. 123, 1269-1277, 1993
 A:Title: Structure and expression of fibulin-2, a novel extracellular matrix protein
 A:Reference number: A49457; MUID:94064787
 A:Accession: A49457

A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1221 <PAN>
 A:Cross-references: GB:X75285; NID:g437046; PID:CAA53040.1; PID:g437047
 R:Sasaki, T.; Mann, K.; Murphy, G.; Chu, M.L.; Timpl, R.
 Eur. J. Biochem. 240, 427-434, 1996
 A:Title: Different susceptibilities of fibulin-1 and fibulin-2 to cleavage by matrix
 A:Reference number: S74094; MUID:96439073
 A:Accession: S74095

A:Molecule type: protein
 A:Residues: 236-238, 'X', 240-247; 260-275; 336-344, 'L', 346-361; 405-426; 566-568, 'EM', 569-
 C:Superfamily: unassigned EGF-related proteins; EGF homology
 C:Keywords: calcium binding; duplication; extracellular matrix; glycoprotein; homotri
 F:942-978/Domain: EGF homology <EGF>

Query Match 66.1%; Score 39; DB 2; Length 1221;
 Best Local Similarity 85.7%; Pred. No. 1.5e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 SCPPGAG 9
 |||||
 Db 98 SCPPGGG 104

RESULT 9
 A54774
 ATP binding cassette transporter ABC1 - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 05-Apr-1995 #sequence_revision 05-Apr-1995 #text_change 17-Mar-2000
 C:Accession: A54774
 R:Luciani, M.F.; Denizot, F.; Savary, S.; Mattei, M.G.; Chlmini, G.
 Genomics 21, 150-159, 1994
 A:Title: Cloning of two novel ABC transporters mapping on human chromosome 9.
 A:Reference number: A54774; MUID:94375008
 A:Accession: A54774
 A:Molecule type: mRNA
 A:Residues: 1-2201 <LUC>

A:Cross-references: GB:X75926; NID:g495256; PIDN:CAA53530.1; PID:g495257
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
C:Keywords: ATP; duplication; P-loop
F:856-1047/Domain: ATP-binding cassette homology <ABC1>
F:873-880/Region: nucleotide-binding motif A (P-loop)
F:1869-2060/Domain: ATP-binding cassette homology <ABC2>
F:1886-1893/Region: nucleotide-binding motif A (P-loop)

Query Match 66.1%; Score 39; DB 2; Length 2201;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 CPPGAG 9
| | | | |
Db 1417 CPPGAG 1422

RESULT 10
B70765
probable cobM - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: B70765
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A:Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987
A:Accession: B70765
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-251 <COL>
A:Cross-references: GB:Z73966; GB:AL123456; NID:g3261577; PIDN:CAA98218.1; PID:g1370237
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: cobM
C:Superfamily: precorrin-3 methylase

Query Match 64.4%; Score 38; DB 2; Length 251;
Best Local Similarity 66.7%; Pred. No. 53;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ASCPPGAGV 10
| | | | |
Db 43 AOCPPGATI 51

RESULT 11
F64084
3-deoxy-manno-octulosonate cytidyltransferase (EC 2.7.7.38) [similarity] - Haemophilus
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 18-Aug-2000
C:Accession: F64084
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M. Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630
A:Accession: F64084
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-427 <TIGR>
A:Cross-references: GB:U32748; GB:I42023; NID:g1573645; PIDN:AAC22311.1; PID:g1573651; T
C:Function:
A:Description: catalyzes the transfer of 3-deoxy-D-manno-octulosonic acid from its CMP c
C:Superfamily: Chlamydia trachomatis 3-deoxy-manno-octulosonate cytidyltransferase
C:Keywords: nucleotidyltransferase; transmembrane protein

Query Match 64.4%; Score 38; DB 2; Length 427;
Best Local Similarity 60.0%; Pred. No. 84;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 HASCPPGAGV 10
| | | | |
Db 44 NASCPGPGI 53

RESULT 12
H70854
hypothetical protein RV2997 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: H70854
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A:Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:98295987
A:Accession: H70854
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-480 <COL>
A:Cross-references: GB:AL021287; GB:AL123456; NID:g3261508; PIDN:CAA16082.1; PID:g279
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV2997
C:Superfamily: phytoene dehydrogenase

Query Match 64.4%; Score 38; DB 2; Length 480;
Best Local Similarity 77.8%; Pred. No. 93;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ASCPPGAGV 10
| | | | |
Db 441 AATPPGAGV 449

RESULT 13
S66498
M-sema F protein precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 24-Nov-1999
C:Accession: S66498
R:Inagaki, S.; Furuyama, T.; Iwahashi, Y. FEBS Lett. 370, 269-272, 1995
A:Title: Identification of a member of mouse semaphorin family.
A:Reference number: S66498; MUID:95385809
A:Accession: S66498
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-834 <INA>
A:Cross-references: EMBL:S79463; NID:g1110598; PIDN:AAB35184.1; PID:g1110599
C:Superfamily: semaphorin
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-834/Product: M-sema F protein #status predicted <MAT>

Query Match 64.4%; Score 38; DB 2; Length 834;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 HASCPPGAG 9
| | | | |
Db 750 HARCQPGG 758

RESULT 14
QRHUP
progesterone receptor form B - human
N:Alternate names: hPR
N:Contains: progesterone receptor form A
C:Species: Homo sapiens (man)
C:Date: 30-Jun-1987 #sequence_revision 18-Nov-1994 #text_change 22-Jun-1999
C:Accession: S09971; S12464; A03245
R:Kastner, P.; Krust, A.; Turcotte, B.; Stropp, U.; Tora, L.; Gronemeyer, H.; Chambon, P.
EMBO J. 9, 1603-1614, 1990
A:Title: Two distinct estrogen-regulated promoters generate transcripts encoding the two
A:Reference number: S09971; MUID:90228361
A:Accession: S09971
A:Molecule type: mRNA
A:Residues: 1-933 <KAS>
A:Cross-references: EMBL:X51730
R:Kastner, P.
submitted to the EMBL Data Library, February 1990
A:Reference number: S12464
A:Accession: S12464
A:Molecule type: mRNA
A:Residues: 1-343, 'T', 345-933 <KA2>
A:Cross-references: EMBL:X51730; NID:g35651; PIDN:CAA36018.1; PID:g35652
R:Mistrali, M.; Atger, M.; D'Auriol, L.; Loosfelt, H.; Meriel, C.; Fridlansky, F.; Guioch
Biochem. Biophys. Res. Commun. 143, 740-748, 1987
A:Title: Complete amino acid sequence of the human progesterone receptor deduced from c
A:Reference number: A03245; MUID:87184565
A:Accession: A03245
A:Molecule type: mRNA
A:Residues: 1-225, 'G', 227-255, 'V', 257-659, 'V', 661-933 <MIS>
A:Cross-references: GB:M15716; NID:g189934; PIDN:AAA60081.1; PID:g189935
C:Genetics:
A:Gene: GDB:PCR
A:Cross-references: GDB:119493; OMIM:264080
A:Map position: 11q22.1-11q22.3
C:Superfamily: progesterone receptor; erba transforming protein homology
C:Keywords: alternative splicing; DNA binding; nucleus; phosphoprotein; steroid hormone
F:1-933/Product: progesterone receptor form B #status predicted <NA1>
F:165-933/Product: progesterone receptor form A #status predicted <MA2>
F:565-829/Domain: erba transforming protein homology <ERBA>
F:567-587/Region: zinc finger CCCC motif
F:603-627/Region: zinc finger CCCC motif
F:681-933/Domain: steroid binding #status predicted <STB>
F:41/Binding site: phosphate (Thr) (covalent) #status predicted
F:227,232,552,793/Binding site: phosphate (Ser) (covalent) #status predicted
F:329,374,601/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 64.4%; Score 38; DB 1; Length 933;
Best Local Similarity 85.7%; Pred. NO. 1.7e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ASCPPGA 8
1:|||||
Db 254 AACPPGA 260

RESULT 15
T46330
hypotheical protein DKFzp434D0513.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C:Accession: T46330
R:Koehler, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wilmann, S.
submitted to the Protein Sequence Database, January 2000
A:Reference number: 223037
A:Accession: T46330
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-981 <AAA>
A:Cross-references: EMBL:AL137259
A:Experimental source: adult testis; clone DKFzp434D0513
C:Genetics:

A:Note: DKFzp434D0513.1

Query Match 64.4%; Score 38; DB 2; Length 981;
Best Local Similarity 85.7%; Pred. NO. 1.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 SCPPGAG 9
1:|||||
Db 959 SCPPGEG 965

Search completed: January 30, 2001, 16:50:39
Job time: 729 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 17:06:12 ; Search time 80.04 seconds
(without alignments)
4.035 Million cell updates/sec

Title: US-09-518-931-4_COPY_129_138
Perfect score: 59
Sequence: 1 HASCPPGAGV 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	40	67.8	277	1 CD40_HUMAN	P25942 homo sapien
2	40	67.8	2813	1 VWF_CANFA	P28295 canis famil
3	40	67.8	2813	1 VWF_HUMAN	P04275 homo sapien
4	39	66.1	1184	1 FBL2_HUMAN	P98095 homo sapien
5	39	66.1	1221	1 FBL2_MOUSE	P37889 mus musculu
6	39	66.1	2201	1 ABC1_HUMAN	O95477 homo sapien
7	39	66.1	2201	1 ABC1_MOUSE	P41233 mus musculu
8	38	64.4	251	1 COBM_MYCTU	Q10672 mycobacteri
9	38	64.4	427	1 KDTA_HAEIN	P44806 haemophilus
10	38	64.4	685	1 TRFE_PAROL	O93429 paralichthy
11	38	64.4	834	1 SM4C_MOUSE	Q64151 mus musculu
12	38	64.4	1964	1 NTC4_MOUSE	P31695 mus musculu
13	38	64.4	3635	1 LMA5_MOUSE	Q61001 mus musculu
14	37	62.7	332	1 NPDI_MOUSE	Q64322 mus musculu
15	37	62.7	368	1 LNK_RAT	P50745 rattus norv
16	37	62.7	392	1 GAG_BLVAV	P25058 bovine leuk
17	37	62.7	547	1 FUMA_ECOLI	P00923 escherichia
18	37	62.7	548	1 FUMB_ECOLI	P14407 escherichia
19	37	62.7	585	1 CIKD_RAT	P25122 rattus norv
20	37	62.7	994	1 MERK_MOUSE	Q60805 mus musculu
21	37	62.7	1001	1 TP3A_HUMAN	Q13472 homo sapien
22	37	62.7	2224	1 FA5_HUMAN	P12259 homo sapien
23	36	61.0	754	1 ECEL_CAVPO	P97739 cavia porce
24	36	61.0	2318	1 NTC3_MOUSE	Q61982 mus musculu
25	36	61.0	2491	1 MPRI_HUMAN	P11717 homo sapien
26	35	59.3	187	1 DHML_METME	Q59543 methylophil
27	35	59.3	321	1 TA29_TOBAC	P24804 nicotiana t
28	35	59.3	335	1 FASA_HUMAN	P25445 homo sapien
29	35	59.3	423	1 POLS_SINDW	P11259 sindbis vir
30	35	59.3	504	1 TRFS_HUMAN	Q13568 homo sapien
31	35	59.3	517	1 FGR_MOUSE	P14234 mus musculu
32	35	59.3	529	1 FGR_HUMAN	P09769 homo sapien
33	35	59.3	538	1 REL_MELGA	P01125 meleagris g

RESULT 1

ID	CD40_HUMAN	STANDARD;	PRT;	277 AA.
AC	P25942;			
DT	01-MAY-1992 (Rel. 22, Created)			
DT	01-MAY-1992 (Rel. 22, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	CD40L RECEPTOR PRECURSOR (B-CELL SURFACE ANTIGEN CD40) (BP50) (CDW40)			
DE	(TUMOR NECROSIS FACTOR RECEPTOR 5).			
GN	TNFRSF5 OR CD40.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89356608; PubMed=2475341;			
RA	Stamenkovic I., Clark E.A., Seed B.;			
RT	"A B-lymphocyte activation molecule related to the nerve growth			
RT	factor receptor and induced by cytokines in carcinomas.";			
RL	EMBO J. 8:1403-1410(1989).			
RN	[2]			
RP	3D-STRUCTURE MODELING OF 24-144.			
RX	MEDLINE=97189482; PubMed=9037712;			
RA	Bajorath J., Aruffo A.;			
RT	"Construction and analysis of a detailed three-dimensional model of			
RT	the ligand binding domain of the human B cell receptor CD40.";			
RL	Proteins 27:59-70(1997).			
RN	[3]			
RP	3D-STRUCTURE MODELING OF 26-186 IN COMPLEX WITH CD40L.			
RX	MEDLINE=98266353; PubMed=9605317;			
RA	Singh J., Garber E., van Vlijmen H., Karpusas M., Hsu Y.-M.,			
RA	Zheng Z., Naismith J.H., Thomas D.;			
RT	"The role of polar interactions in the molecular recognition of CD40L			
RT	with its receptor CD40.";			
RL	Protein Sci. 7:1124-1135(1998).			
CC	-!- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD40L.			
CC	-!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.			
CC	-!- TISSUE SPECIFICITY: B-CELLS AND IN PRIMARY CARCINOMAS.			
CC	-!- SIMILARITY: CONTAINS A LA-NGR/TNFR-TYPE CYSTEINE-RICH REGION.			
CC	-!- DATABASE: NAME=PROW; NOTE=CD guide CD40 entry;			
CC	WWW='http://www.ncbi.nlm.nih.gov/prow/cd/cd40.htm'.			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: X60592; CAA43045.1; -
DR PIR: S04460; S04460.
DR PDB: ICDF; 01-APR-97.
DR MIN; 109535; -
DR INTERPRO; IPR001368; -
DR PFAM; PF00020; TNFR_c6; 4.

34	35	59.3	545	1	FGR_FSVGR	P00544 feline sarc
35	35	59.3	567	1	FACC_BOVIN	O19104 bos taurus
36	35	59.3	754	1	ECEL_BOVIN	P42891 bos taurus
37	35	59.3	770	1	ECEL_HUMAN	P42892 homo sapien
38	35	59.3	1016	1	FDOG_ECOLI	P32176 escherichia
39	35	59.3	1236	1	POLS_WEEV	P13897 western equ
40	35	59.3	1245	1	POLS_SINDO	P27285 sindbis vir
41	35	59.3	1245	1	POLS_SINDV	P27285 sindbis vir
42	35	59.3	2182	1	POLG_CXB1J	P08291 c genome po
43	35	59.3	2191	1	POLG_EC06C	Q66474 e genome po
44	35	59.3	2476	1	ZAN_PIG	Q28983 sus scrofa
45	34.5	58.5	1305	1	GAK_RAT	P97874 rattus norv

ALIGNMENTS

FT CARBOHYD 2290 2290 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2357 2357 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2400 2400 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2546 2546 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2585 2585 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2790 2790 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SITE 531 533 CELL ATTACHMENT SITE (POTENTIAL).
 FT SITE 698 700 CELL ATTACHMENT SITE (POTENTIAL).
 FT SITE 2507 2509 CELL ATTACHMENT SITE (POTENTIAL).
 SQ SEQUENCE 2813 AA; 309730 MW; 03D28D64577188D5 CRC64;

Query Match 67.8%; Score 40; DB 1; Length 2813;
 Best Local Similarity 85.7%; Pred. No. 97;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 HASCPPG 7
 II IIII
 Db 1174 HAHCPPG 1180

RESULT 3
 VWF_HUMAN
 ID VWF_HUMAN STANDARD; PRT; 2813 AA.
 AC P04275;
 DT 20-MAR-1987 (Rel. 04, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE VON WILLEBRAND FACTOR PRECURSOR (VWF).
 GN FBVWF OR VWF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90062044; PubMed=2584182;
 RA Mancuso D.J., Tuley E.A., Westfield L.A., Worrall N.K.,
 RA Shelton-Inloes B.B., Sorace J.M., Alevis Y.G., Sadler J.E.;
 RT "Structure of the gene for human von Willebrand factor.";
 RL J. Biol. Chem. 264:19514-19527(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87016349; PubMed=3489923;
 RA Bonthron D., Orr E.C., Mitscock L.M., Ginsburg D., Handin R.I.,
 RA Orkin S.H.;
 RT "Nucleotide sequence of pre-pro-von Willebrand factor cDNA.";
 RL Nucleic Acids Res. 14:7125-7128(1986).
 RN [3]
 RP SEQUENCE OF 1-120 FROM N.A., AND SEQUENCE OF 23-56.
 RC TISSUE=UMBILICAL VEIN ENDOTHELIAL CELLS;
 RX MEDLINE=87213253; PubMed=3495266;
 RA Shelton-Inloes B.B., Broze G.J. Jr., Miletich J.P., Sadler J.E.;
 RT "Evolution of human von Willebrand factor: cDNA sequence
 RT polymorphisms, repeated domains, and relationship to von Willebrand
 RT antigen II.";
 RL Biochem. Biophys. Res. Commun. 144:657-665(1987).
 RN [4]
 RP SEQUENCE OF 1-1400 FROM N.A.
 RX MEDLINE=87004550; PubMed=3019665;
 RA Verweij C.L., Diergaarde P.J., Hart M., Pannekoek H.;
 RT "Full-length von Willebrand factor (vWF) cDNA encodes a highly
 RT repetitive protein considerably larger than the mature vWF subunit.";
 RL EMBO J. 5:1839-1847(1986).
 RN [5]
 RP ERRATUM.
 RA Verweij C.L., Diergaarde P.J., Hart M., Pannekoek H.;
 RL EMBO J. 5:3074-3074(1986).
 RN [6]
 RP SEQUENCE OF 764-2813.
 RX MEDLINE=86269895; PubMed=3524673;
 RA Titani K., Kumar S., Takio K., Ericsson L.H., Wade R.D., Ashida K.,
 RA Walsh K.A., Chopek M.W., Sadler J.E., Fujikawa K.;
 RT "Amino acid sequence of human von Willebrand factor.";

RL Biochemistry 25:3171-3184(1986).
 RN [7]
 RP SEQUENCE OF 781-1424 FROM N.A.
 RX MEDLINE=86269894; PubMed=3488076;
 RA Shelton-Inloes B.B., Titani K., Sadler J.E.;
 RT "cDNA sequences for human von Willebrand factor reveal five types of
 RT repeated domains and five possible protein sequence polymorphisms.";
 RL Biochemistry 25:3164-3171(1986).
 RN [8]
 RP SEQUENCE OF 764-873 AND 1289-2813 FROM N.A.
 RX MEDLINE=86016708; PubMed=2864686;
 RA Sadler J.E., Shelton-Inloes B.B., Sorace J.M., Harlan J.M.,
 RA Titani K., Davie E.W.;
 RT "Cloning and characterization of two cDNAs coding for human von
 RT Willebrand factor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:6394-6398(1985).
 RN [9]
 RP SEQUENCE OF 990-1947 FROM N.A.
 RX MEDLINE=91105089; PubMed=1988024;
 RA Mancuso D.J., Tuley E.A., Westfield L.A., Lester-Mancuso T.L.,
 RA Le Beau M.M., Sorace J.M., Sadler J.E.;
 RT "Human von Willebrand factor gene and pseudogene: structural analysis
 RT and differentiation by polymerase chain reaction.";
 RL Biochemistry 30:253-269(1991).
 RN [10]
 RP SEQUENCE OF 2731-2813 FROM N.A.
 RX MEDLINE=85269603; PubMed=3875078;
 RA Verweij C.L., de Vries C.J.M., Distel B., van Zonneveld A.-J.,
 RA Geurts van Kessel A., van Mourik J.A., Pannekoek H.;
 RT "Construction of cDNA coding for human von Willebrand factor using
 RT antibody probes for colony-screening and mapping of the chromosomal
 RL gene.";
 RN Nucleic Acids Res. 13:4699-4717(1985).
 RN [11]
 RP SEQUENCE OF 1-177 FROM N.A.
 RX MEDLINE=88111704; PubMed=2828057;
 RA Bonthron D., Orkin S.H.;
 RT "The human von Willebrand factor gene. Structure of the 5' region.";
 RL Eur. J. Biochem. 171:51-57(1988).
 RN [12]
 RP SEQUENCE OF 2621-2813 FROM N.A.
 RX MEDLINE=85244588; PubMed=3874428;
 RA Ginsburg D., Handin R.I., Bonthron D.T., Donlon T.A., Bruns G.A.P.,
 RA Latt S.A., Orkin S.H.;
 RT "Human von Willebrand factor (vWF): isolation of complementary DNA
 RT (cDNA) clones and chromosomal localization.";
 RL Science 228:1401-1406(1985).
 RN [13]
 RP SEQUENCE OF 2731-2813 FROM N.A.
 RX MEDLINE=85201687; PubMed=3873280;
 RA Lynch D.C., Zimmerman T.S., Collins C.J., Brown M., Morin M.J.,
 RA Ling E.H., Livingston D.M.;
 RT "Molecular cloning of cDNA for human von Willebrand factor:
 RT authentication by a new method.";
 RL Cell 41:49-56(1985).
 RN [14]
 RP REVISIONS.
 RA Lynch D.C.;
 RL Submitted (JUL-1991) to the EMBL/GenBank/DBJ databases.
 RN [15]
 RP SEQUENCE OF 2731-2813 FROM N.A.
 RX MEDLINE=87260814; PubMed=3496594;
 RA Collins C.J., Underdahl J.P., Levene R.B., Ravera C.P.,
 RA Morin M.J., Dombalagian M.J., Ricca G., Livingston D.M.,
 RA Lynch D.C.;
 RT "Molecular cloning of the human gene for von Willebrand factor and
 RT identification of the transcription initiation site.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:4393-4397(1987).
 RN [16]
 RP DISULFIDE BONDS.
 RX MEDLINE=88163465; PubMed=3502076;
 RA Marti T., Rosselet S.J., Titani K., Walsh K.A.;
 RT "Identification of disulfide-bridged substructures within human von

RT Willebrand factor.;
 RL Biochemistry 26:8099-8109(1987).
 [17]
 RP STRUCTURE OF CARBOHYDRATES.
 RX MEDLINE=86274702; PubMed=3089784;
 RA Samor B., Michalski J.C., Debray H., Mazurier C., Goudemand M.,
 van Halbeek H., Villegienchart J.F.G., Montreuil J.;
 RT "Primary structure of a new tetraantennary glycan of the N-
 acetylglucosaminic type isolated from human factor VIII/von
 Willebrand factor.";
 RL Eur. J. Biochem. 158:295-298(1986).
 [18]
 RN X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 1261-1468.
 RX MEDLINE=98221174; PubMed=9553097;
 RA Emsley J., Cruz M., Handin R., Liddington R.;
 RT "Crystal structure of the von Willebrand factor A1 domain and
 RT implications for the binding of platelet glycoprotein Ib.";
 RL J. Biol. Chem. 273:10396-10401(1998).
 [19]
 RN X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 1685-1873.
 RX MEDLINE=97472999; PubMed=9331419;
 RA Huizinga E.G., Martijn van der Plas R., Kroon J., Sixma J.J., Gros P.;
 RT "Crystal structure of the A3 domain of human von Willebrand factor:
 RT implications for collagen binding.";
 RL Structure 5:1147-1156(1997).
 [20]
 RN X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1686-1872.
 RX MEDLINE=97460108; PubMed=9312128;
 RA Bienkowska J., Cruz M., Atlamo A., Handin R., Liddington R.;
 RT "The von Willebrand factor A3 domain does not contain a metal ion-
 RT dependent adhesion site motif.";
 RL J. Biol. Chem. 272:25162-25167(1997).
 [21]
 RN VARIANTS TRP-1597 AND ASP-1607.
 RX MEDLINE=89264495; PubMed=2786201;
 RA Ginsburg D., Konkli B.A., Gill J.C., Montgomery R.R.,
 Bockenstedt P.L., Johnson T.A., Yang A.Y.;
 RT "Molecular basis of human von Willebrand disease: analysis of
 RT platelet von Willebrand factor mRNA.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:3723-3727(1989).
 [22]
 RN VARIANT THR-1628.
 RX MEDLINE=91196734; PubMed=1673047;
 RA Iannuzzi M.C., Hidaka N., Boehnke M., Bruck M.E., Hanna W.T.,
 Collins F.S., Ginsburg D.;
 RT "Analysis of the relationship of von Willebrand disease (vWD) and
 RT hereditary hemorrhagic telangiectasia and identification of a
 RT potential type IIA vWD mutation (Ile865 to Thr).";
 RL Am. J. Hum. Genet. 48:757-763(1991).
 [23]
 RN VARIANTS NORMANDY-2 AND NORMANDY-3.
 RX MEDLINE=92001464; PubMed=1832934;
 RA Gaucher C., Mercier B., Jorleux S., Oufkir D., Mazurier C.;
 RT "Identification of two point mutations in the von Willebrand factor
 RT gene of three families with the 'Normandy' variant of von Willebrand
 RT disease.";
 RL Br. J. Haematol. 78:506-514(1991).
 [24]
 RN VARIANT CYS-1308.
 RX MEDLINE=92104315; PubMed=1761120;
 RA Donner M., Andersson A.-M., Kristofferson A.-C., Nilsson I.M.,
 Dahlback B., Holmberg L.;
 RT "An Arg545->Cys545 substitution mutation of the von Willebrand
 RT factor in type IIB von Willebrand's disease.";
 RL Eur. J. Haematol. 47:342-345(1991).
 [25]
 RN VARIANTS TRP-1306; CYS-1308 AND PRO-1613.
 RX MEDLINE=91185601; PubMed=2010538;
 RA Randi A.M., Rabinowitz I., Mancuso D.J., Mannucci P.M., Sadler J.E.;
 RT "Molecular basis of von Willebrand disease type IIB. Candidate
 RT mutations cluster in one disulfide loop between proposed platelet
 RT glycoprotein Ib binding sequences.";
 RL J. Clin. Invest. 87:1220-1226(1991).

[26]
 RN VARIANTS TRP-1306; CYS-1308; MET-1316; GLN-1341 AND HIS-1399.
 RP Biochemistry 26:8099-8109(1987).
 RX MEDLINE=91185602; PubMed=1672694;
 RA Cooney K.A., Nichols W.C., Bruck M.E., Bahou W.F., Shapiro A.D.,
 Query Match 67.8%; Score 40; DB 1; Length 2813;
 Best Local Similarity 85.7%; Pred. No. 97;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 HASCPPG 7
 Db 1174 HASCPPG 1180

RESULT 4
 FBL2_HUMAN
 ID FBL2_HUMAN STANDARD; PRT; 1184 AA.
 AC P98095;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE FIBULIN-2 PRECURSOR.
 GN FBLN2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE=FIBROBLAST;
 RX MEDLINE=95104855; PubMed=7806230;
 RA Zhang R.-Z., Pan T.-C., Zhang Z.-Y., Mattei M.-G., Timpl R.,
 Chu M.-L.;
 RT "Fibulin-2 (FBLN2): human cDNA sequence, mRNA expression, and mapping
 RT of the gene on human and mouse chromosomes.";
 RL Genomics 22:425-430(1994).
 CC -!- FUNCTION: ITS BINDING TO FIBRONECTIN AND SOME OTHER LIGANDS IS
 CC CALCIUM DEPENDENT.
 CC -!- SUBUNIT: DISULFIDE-LINKED HOMOTRIMER.
 CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX. COMPONENT OF BOTH
 CC BASEMENT MEMBRANES AND OTHER CONNECTIVE TISSUES.
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN HEART, PLACENTA AND OVARY.
 CC -!- SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.
 CC -!- SIMILARITY: CONTAINS 11 EGF-LIKE DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 EMBL; X82494; CAA57876.1;
 HSP; P07204; IFGD.
 MIM: 135821;
 INTERPRO: IPR000020;
 INTERPRO: IPR000152;
 INTERPRO: IPR000561;
 INTERPRO: IPR001881;
 PFAM: PF01821; ANATO; 2.
 PFAM: PF00008; EGF; 7.
 PROSITE: PS00010; ASX_HYDROXYL; 5.
 PROSITE: PS01177; ANAPHYLATOXIN_1; 3.
 PROSITE: PS01178; ANAPHYLATOXIN_2; 3.
 PROSITE: PS00022; EGF_1; FALSE_NEG.
 PROSITE: PS01186; EGF_2; 5.
 PROSITE: PS01187; EGF_CA; 9.
 Signal: Glycoprotein; Extracellular matrix; Plasma; EGF-like domain;
 Calcium-binding; Repeat.
 SIGNAL 1 27 POTENTIAL.
 FT CHAIN 28 1184 FIBULIN-2.
 FT DOMAIN 28 444 N.
 FT DOMAIN 28 177 SUBDOMAIN NA (CYS-RICH).

FT	DOMAIN	178	444	SUBDOMAIN NB (CYS-FREE).	AC	P37889; Q9WUI2;
FT	DOMAIN	445	553	3 X ANAPHYLATOXIN REPEATS (DOMAIN I).	DT	01-OCT-1994 (Rel. 30, Created)
FT	REPEAT	445	580	ANAPHYLATOXIN-LIKE 1.	DT	01-OCT-1994 (Rel. 30, Last sequence update)
FT	REPEAT	488	519	ANAPHYLATOXIN-LIKE 2.	DT	01-OCT-2000 (Rel. 40, Last annotation update)
FT	REPEAT	521	553	ANAPHYLATOXIN-LIKE 3.	DE	FIBULIN-2 PRECURSOR.
FT	DOMAIN	604	1069	10 X EGF-LIKE REPEATS (DOMAIN II).	GN	FBLN2.
FT	DOMAIN	604	645	EGF-LIKE 1, CALCIUM-BINDING.	OS	Mus musculus (Mouse).
FT	DOMAIN	679	718	EGF-LIKE 2.	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
FT	DOMAIN	719	763	EGF-LIKE 3, CALCIUM-BINDING.	OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
FT	DOMAIN	764	809	EGF-LIKE 4, CALCIUM-BINDING.	RN	[1]
FT	DOMAIN	810	857	EGF-LIKE 5, CALCIUM-BINDING.	RP	SEQUENCE FROM N.A., AND SEQUENCE OF 27-35.
FT	DOMAIN	858	900	EGF-LIKE 6, CALCIUM-BINDING.	RC	TISSUE=FIBROBLAST;
FT	DOMAIN	901	942	EGF-LIKE 7, CALCIUM-BINDING.	RX	MEDLINE=94064787; PubMed=8245130;
FT	DOMAIN	942	981	EGF-LIKE 8, CALCIUM-BINDING.	RA	Pan T.-C., Sasaki T., Zhang R.-Z., Faessler R., Timpl R., Chu M.-L.;
FT	DOMAIN	983	1024	EGF-LIKE 9, CALCIUM-BINDING.	RT	"Structure and expression of fibulin-2, a novel extracellular matrix
FT	DOMAIN	1025	1069	EGF-LIKE 10, CALCIUM-BINDING.	RT	protein with multiple EGF-like repeats and consensus motifs for
FT	DOMAIN	1070	1184	DOMAIN III.	RL	calcium binding.";
FT	DISULFID	445	472	BY SIMILARITY.	RL	J. Cell Biol. 123:1269-1277(1993).
FT	DISULFID	446	479	BY SIMILARITY.	RN	[2]
FT	DISULFID	459	480	BY SIMILARITY.	RP	SEQUENCE FROM N.A.
FT	DISULFID	489	518	BY SIMILARITY.	RX	MEDLINE=99337686; PubMed=10406956;
FT	DISULFID	502	519	BY SIMILARITY.	RA	Graessel S., Sicot F.-X., Gotta S., Chu M.-L.;
FT	DISULFID	521	545	BY SIMILARITY.	RT	"Mouse fibulin-2 gene. Complete exon-intron organization and promoter
FT	DISULFID	522	552	BY SIMILARITY.	RT	characterization.";
FT	DISULFID	535	553	BY SIMILARITY.	RL	Eur. J. Biochem. 263:471-477(1999).
FT	DISULFID	608	620	BY SIMILARITY.	CC	-1- FUNCTION: ITS BINDING TO FIBRONECTIN AND SOME OTHER LIGANDS IS
FT	DISULFID	616	629	BY SIMILARITY.	CC	CALCIUM DEPENDENT.
FT	DISULFID	631	644	BY SIMILARITY.	CC	-1- SUBUNIT: DISULFIDE-LINKED HOMOTRIMER.
FT	DISULFID	683	693	BY SIMILARITY.	CC	-1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.
FT	DISULFID	689	702	BY SIMILARITY.	CC	-1- ALTERNATIVE PRODUCTS: AT LEAST TWO FORMS OF FIBULIN-2 DUE TO THE
FT	DISULFID	704	717	BY SIMILARITY.	CC	PRESENCE OR ABSENCE OF A SINGLE EGF-LIKE (3) REPEAT MAY ARISE BY
FT	DISULFID	723	736	BY SIMILARITY.	CC	ALTERNATIVE SPLICING.
FT	DISULFID	730	745	BY SIMILARITY.	CC	-1- TISSUE SPECIFICITY: COMPONENT OF BOTH BASEMENT MEMBRANES AND OTHER
FT	DISULFID	751	762	BY SIMILARITY.	CC	CONNECTIVE TISSUES.
FT	DISULFID	768	781	BY SIMILARITY.	CC	-1- SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.
FT	DISULFID	775	790	BY SIMILARITY.	CC	-1- SIMILARITY: CONTAINS 11 EGF-LIKE DOMAINS.
FT	DISULFID	796	808	BY SIMILARITY.	CC	-----
FT	DISULFID	814	827	BY SIMILARITY.	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
FT	DISULFID	821	836	BY SIMILARITY.	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
FT	DISULFID	843	856	BY SIMILARITY.	CC	the European Bioinformatics Institute. There are no restrictions on its
FT	DISULFID	862	875	BY SIMILARITY.	CC	use by non-profit institutions as long as its content is in no way
FT	DISULFID	869	884	BY SIMILARITY.	CC	modified and this statement is not removed. Usage by and for commercial
FT	DISULFID	886	899	BY SIMILARITY.	CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
FT	DISULFID	905	917	BY SIMILARITY.	CC	or send an email to license@isb-sib.ch).
FT	DISULFID	913	926	BY SIMILARITY.	CC	-----
FT	DISULFID	928	941	BY SIMILARITY.	CC	EMBL; X75285; CAA53040.1; -.
FT	DISULFID	947	956	BY SIMILARITY.	DR	EMBL; AF135253; AAD34456.1; -.
FT	DISULFID	952	965	BY SIMILARITY.	DR	EMBL; AF135239; AAD34456.1; JOINED.
FT	DISULFID	967	980	BY SIMILARITY.	DR	EMBL; AF135240; AAD34456.1; JOINED.
FT	DISULFID	986	998	BY SIMILARITY.	DR	EMBL; AF135241; AAD34456.1; JOINED.
FT	DISULFID	994	1007	BY SIMILARITY.	DR	EMBL; AF135242; AAD34456.1; JOINED.
FT	DISULFID	1009	1023	BY SIMILARITY.	DR	EMBL; AF135243; AAD34456.1; JOINED.
FT	DISULFID	1029	1042	BY SIMILARITY.	DR	EMBL; AF135244; AAD34456.1; JOINED.
FT	DISULFID	1036	1051	BY SIMILARITY.	DR	EMBL; AF135245; AAD34456.1; JOINED.
FT	DISULFID	1056	1068	BY SIMILARITY.	DR	EMBL; AF135246; AAD34456.1; JOINED.
FT	CARBOHYD	180	180	N-LINKED (GLCNAC. . .) (POTENTIAL).	DR	EMBL; AF135247; AAD34456.1; JOINED.
FT	CARBOHYD	507	507	N-LINKED (GLCNAC. . .) (POTENTIAL).	DR	EMBL; AF135248; AAD34456.1; JOINED.
FT	CARBOHYD	507	507	N-LINKED (GLCNAC. . .) (POTENTIAL).	DR	EMBL; AF135249; AAD34456.1; JOINED.
FT	CARBOHYD	1035	1035	N-LINKED (GLCNAC. . .) (POTENTIAL).	DR	EMBL; AF135250; AAD34456.1; JOINED.
FT	CARBOHYD	1035	1035	N-LINKED (GLCNAC. . .) (POTENTIAL).	DR	EMBL; AF135251; AAD34456.1; JOINED.
FT	CARBOHYD	1035	1035	N-LINKED (GLCNAC. . .) (POTENTIAL).	DR	EMBL; AF135252; AAD34456.1; JOINED.
FT	CARBOHYD	1035	1035	N-LINKED (GLCNAC. . .) (POTENTIAL).	DR	PIR; A49457; A49457.
FT	CARBOHYD	1035	1035	N-LINKED (GLCNAC. . .) (POTENTIAL).	DR	HSP; P07204; IFGD.
FT	CARBOHYD	1035	1035	N-LINKED (GLCNAC. . .) (POTENTIAL).	DR	MGD; MGI:95488; FBLN2.
FT	CARBOHYD	1035	1035	N-LINKED (GLCNAC. . .) (POTENTIAL).	DR	INTERPRO: IPR000020; -.
FT	CARBOHYD	1035	1035	N-LINKED (GLCNAC. . .) (POTENTIAL).	DR	INTERPRO: IPR000152; -.
FT	CARBOHYD	1035	1035	N-LINKED (GLCNAC. . .) (POTENTIAL).	DR	INTERPRO: IPR000561; -.
FT	CARBOHYD	1035	1035	N-LINKED (GLCNAC. . .) (POTENTIAL).	DR	INTERPRO: IPR001881; -.
FT	CARBOHYD	1035	1035	N-LINKED (GLCNAC. . .) (POTENTIAL).	DR	PFAM; PF01821; ANATO; 2.
FT	CARBOHYD	1035	1035	N-LINKED (GLCNAC. . .) (POTENTIAL).	DR	PFAM; PF00008; EGF; 6.
FT	CARBOHYD	1035	1035	N-LINKED (GLCNAC. . .) (POTENTIAL).	DR	PROSITE; PS00010; ASX_HYDROXYL; 5.
FT	CARBOHYD	1035	1035	N-LINKED (GLCNAC. . .) (POTENTIAL).	DR	PROSITE; PS01177; ANAPHYLATOXIN_1; 3.
FT	CARBOHYD	1035	1035	N-LINKED (GLCNAC. . .) (POTENTIAL).	DR	PROSITE; PS01178; ANAPHYLATOXIN_2; 3.
FT	CARBOHYD	1035	1035	N-LINKED (GLCNAC. . .) (POTENTIAL).	DR	PROSITE; PS01178; ANAPHYLATOXIN_2; 3.

Query Match 66.1%; Score 39; DB 1; Length 1184;
 Best Local Similarity 85.7%; Pred. No. 64;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Caps 0;

Qy 3 SCPPGAG 9
 |||||
 Db 99 SCPPGGG 105

RESULT 5
 FBL2_MOUSE
 ID FBL2_MOUSE STANDARD; PRT; 1221 AA.

RT "Cloning of two novel ABC transporters mapping on human chromosome
9.";
RL Genomics 21:150-159(1994).
CC -!- FUNCTION: CAMP-DEPENDENT AND SULFONYLUREA-SENSITIVE ANION
CC TRANSPORTER. KEY GATEKEEPER INFLUENCING INTRACELLULAR CHOLESTEROL
CC TRANSPORT (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED IN ADULT TISSUES. HIGHEST
CC LEVELS ARE FOUND IN PREGNANT UTERUS AND UTERUS.
CC -!- DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES,
CC EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN
CC ATP BINDING CASSETTE (ABC) DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS). MDR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X75926; CAA53530.1; -;
CC MGD; MGI:99607; ABCA1.
CC INTERPRO: IPR001617; -;
CC PFAM: PF00005; ABC_tran; 2.
CC PROSITE: PS00211; ABC_TRANSPORTER; 1.
CC ATP-binding: Glycoprotein; Transmembrane; Transport.
CC TRANSMEM 580 596 POTENTIAL.
CC TRANSMEM 630 646 POTENTIAL.
CC TRANSMEM 657 673 POTENTIAL.
CC TRANSMEM 689 705 POTENTIAL.
CC TRANSMEM 711 727 POTENTIAL.
CC TRANSMEM 981 997 POTENTIAL.
CC TRANSMEM 1291 1307 POTENTIAL.
CC TRANSMEM 1601 1617 POTENTIAL.
CC TRANSMEM 1648 1664 POTENTIAL.
CC TRANSMEM 1677 1693 POTENTIAL.
CC TRANSMEM 1715 1731 POTENTIAL.
CC TRANSMEM 1794 1810 POTENTIAL.
CC NP_BIND 873 880 ATP (POTENTIAL).
CC NP_BIND 886 893 ATP (POTENTIAL).
CC CARBOHYD 38 91 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 91 91 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 101 101 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 136 136 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 184 184 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 232 232 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 277 277 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 289 289 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 340 340 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 418 418 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 429 429 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 461 461 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 760 760 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 1084 1084 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 1234 1234 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 1393 1393 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 1439 1439 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 1444 1444 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 1577 1577 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 1984 1984 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 2178 2178 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SEQUENCE 2201 AA; 246686 MW; C7DB9A6CB7A5DB73 CRC64;

Query Match 66.1%; Score 39; DB 1; Length 2201;
Best Local Similarity 100.0%; Pred. NO. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 4 CPPGAG 9
| | | | |
DB 1417 CPPGAG 1422

RESULT 8
COBM_MYCTU
ID COBM_MYCTU STANDARD; PRT; 251 AA.
AC Q10672;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PRECORRIN-4 C11-METHYLTRANSFERASE (EC 2.1.1.133) (PRECORRIN-3
DE METHYLASE).
GN COBM OR RV2071C OR MTCV49.10C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RC MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeller K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
RA Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
CC -!- FUNCTION: METHYLTRANSFERASE THAT CATALYZES THE METHYLATION OF
CC C-11 IN PRECORRIN-4 TO FORM PRECORRIN-5.
CC -!- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + PRECORRIN-4 = S-
CC ADENOSYL-L-HOMOCYSTEINE + PRECORRIN 5.
CC -!- PATHWAY: COBALAMIN BIOSYNTHESIS.
CC -!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS SUMT, CYSG, CBIF/COBM
CC AND CBIL/COBI.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z73966; CAA98218.1; -;
CC TUBERCULIST; RV2071C; -;
CC INTERPRO: IPR000878; -;
CC INTERPRO: IPR003043; -;
CC PFAM: PF00590; TP_methylase; 1.
CC PROSITE: PS00839; SUMT_1; 1.
CC PROSITE: PS00840; SUMT_2; 1.
CC Cobalamin biosynthesis; Porphyrin biosynthesis; Transferase;
KW Methyltransferase.
SQ SEQUENCE 251 AA; 26420 MW; 8DBF213ABFC91CD3 CRC64;

Query Match 64.4%; Score 38; DB 1; Length 251;
Best Local Similarity 66.7%; Pred. NO. 23;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 ASCPPGAGV 10
| | | | |
DB 43 AOCPPGATI 51

RESULT 9
KDTA_HAEIN
ID KDTA_HAEIN STANDARD; PRT; 427 AA.
AC P44806;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE 3-DEOXY-D-MANNO-OCTULOSONIC-ACID TRANSFERASE (EC 2.2.1.11) (KDO
 TRANSFERASE).
 GN KDTA OR WAAA OR HI0652.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellales;
 OC Haemophilus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RD / KW20;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudke D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus
 RT influenzae Rd.";
 RL Science 269:496-512(1995).
 CC -1- FUNCTION: ESSENTIAL STEP IN LIPOPOLYSACCHARIDES BIOSYNTHESIS.
 CC ACTS AT TRANSFER OF 3-DEOXY-D-MONO OCTULONIC ACID (KDO) FROM
 CC CMP-KDO TO A TETRAACYLDISACCHARIDE 1,4'-BISPHOSPHATE PRECURSOR
 CC OF LIPID A (LIPID IVA) (BY SIMILARITY).
 CC -1- PATHWAY: LIPOPOLYSACCHARIDE CORE BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: INNER MEMBRANE-ANCHORED (POSSIBLE).
 CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U32748; AAC2311.1; -;
 DR TIGR: HI0652; -;
 DR INTERPRO: IPR001296; -;
 DR PFAM: PF00534; Glycos_transf_1; 1.
 DR Transferase: Lipopolysaccharide biosynthesis; Transmembrane;
 KW Inner membrane; Signal-anchor.
 FT TRANSMEM 4 24
 FT SIGNAL-ANCHOR (POTENTIAL).
 SQ SEQUENCE 427 AA; 48742 MW; 2B007E5B47EDD17D CRC64;

 Query Match 64.4%; Score 38; DB 1; Length 427;
 Best Local Similarity 60.0%; Pred. No. 37;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

 Qy 1 HASCPPGAGV 10
 Db :||||| I:
 44 NASCPPQGI 53

 RESULT 10
 TRFE_PAROL STANDARD; PRT; 685 AA.
 AC O93429;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE SEROTRANSFERRIN PRECURSOR.
 OS Paralicthys olivaceus (Flounder).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
 OC Pleuronectoidei; Bothidae; Paralicthys.
 RN [1]
 RP SEQUENCE FROM N.A.

 RA Kim Y., Lee J., Hong Y., Hirono I., Aoki T.;
 RT "Molecular cloning and sequence analysis of transferrin cDNA from
 RT Japanese flounder Paralicthys olivaceus.";
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
 CC CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
 CC OF AN ANION, USUALLY BICARBONATE.
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
 CC -1- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D88801; BAA28944.1; -;
 DR INTERPRO: IPR001156; -;
 DR PFAM: PF00405; transferrin; 2.
 DR PRINTS: PR00422; TRANSFERRIN.
 DR PROSITE: PS00205; TRANSFERRIN_1; 2.
 DR PROSITE: PS00206; TRANSFERRIN_2; 2.
 DR PROSITE: PS00207; TRANSFERRIN_3; 1.
 KW Iron transport; Glycoprotein; Metal-binding; Signal.
 FT SIGNAL 1 16
 FT CHAIN 17 685
 FT SEROTRANSFERRIN.
 FT DISULFID 26 48
 FT DISULFID 125 206
 FT DISULFID 170 184
 FT DISULFID 234 248
 FT DISULFID 343 379
 FT DISULFID 353 370
 FT DISULFID 404 678
 FT DISULFID 419 639
 FT DISULFID 451 526
 FT DISULFID 475 667
 FT DISULFID 485 499
 FT DISULFID 496 509
 FT DISULFID 566 580
 FT METAL 72 72
 FT METAL 102 102
 FT METAL 200 200
 FT METAL 256 256
 FT METAL 394 394
 FT METAL 428 428
 FT METAL 520 520
 FT METAL 588 588
 FT BINDING 457 457
 FT CARBOHYD 476 476
 FT N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 685 AA; 74654 MW; 5A6F622AD7D5B662 CRC64;

 Query Match 64.4%; Score 38; DB 1; Length 685;
 Best Local Similarity 75.0%; Pred. No. 57;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

 Qy 1 HASCPPGA 8
 Db :||| |||
 167 HSSCAPGA 174

 RESULT 11
 SM4C_MOUSE STANDARD; PRT; 834 AA.
 AC Q64151;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE SEMAPHORIN 4C PRECURSOR (SEMAPHORIN I) (SEMA I) (SEMAPHORIN C-LIKE 1)

```

DE (M-SEMA F).
GN SEMA4C OR SEMA1 OR SEMA11.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=92194507; PubMed=1312643;
RX Robbins J., Blondel B.J., Gallahan D., Callahan R.;
RT "Mouse mammary tumor gene int-3: a member of the notch gene family
RL transforms mammary epithelial cells.";
RL J. Virol. 66:2594-2599(1992).
RN [2]
RP REVISIONS, SEQUENCE FROM N.A.
RX MEDLINE=97294599; PubMed=9150355;
RA Gallahan D., Callahan R.;
RT "The mouse mammary tumor associated gene INT3 is a unique member of
RL the NOTCH gene family (NOTCH4).";
RN Oncogene 14:1883-1890(1997);
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG, AND TESTIS;
RX MEDLINE=96281668; PubMed=8681805;
RA Uyttendaele H., Marazzi G., Wu G., Yan Q., Sassoon D., Kitajewski J.;
RT "Notch4/int-3, a mammary proto-oncogene, is an endothelial
RL cell-specific mammalian Notch gene.";
RL Development 122:2251-2259(1996).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- DISEASE: ACTIVATED INT-3 TRANSFORMS MAMMARY EPITHELIAL CELLS.
CC -!- SIMILARITY: CONTAINS 29 EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC -!- SIMILARITY: CONTAINS 6 CDC10/SWI6 REPEATS.
CC -!- SIMILARITY: CONTAINS 6 ANK REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M80456; AAB38377.1; -
DR EMBL: U43691; AAC52630.1; -
DR PIR: A38072; TWMVT3.
DR HSP: P00740; IIXA.
DR MGD: MGI:107471; NOTCH4.
DR INTERPRO: IPR00152; -
DR INTERPRO: IPR000561; -
DR INTERPRO: IPR000800; -
DR INTERPRO: IPR001438; -
DR INTERPRO: IPR001881; -
DR INTERPRO: IPR002110; -
DR PFAM: PF00008; EGF; 27.
DR PFAM: PF00023; ank; 6.
DR PFAM: PF00066; notch; 2.
DR PRINTS: PRO0010; EGFLOOD.
DR PROSITE: PS50088; ANK_REPEAT; 5.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE: PS00010; ASX_HYDROXYL; 11.
DR PROSITE: PS00022; EGF_1; 28.
DR PROSITE: PS01186; EGF_2; 21.
DR PROSITE: PS01187; EGF_CA; 9.
DR Differentiation: Neurogenesis; Repeat; EGF-like domain; Transmembrane;
KW Glycoprotein; Proto-oncogene; ANK repeat; Signal.
FT SIGNAL 1 20
FT CHAIN 21 1964
FT DOMAIN 21 1443
FT TRANSMEM 1444 1464
FT DOMAIN 1465 1964
FT DOMAIN 21 60
FT DOMAIN 61 112
FT DOMAIN 115 152
FT DOMAIN 115 152

DE (M-SEMA F).
GN SEMA4C OR SEMA1 OR SEMA11.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=NEONATAL BRAIN;
RX MEDLINE=95385809; PubMed=7656991;
RA Inagaki S., Furuyama T., Iwahashi Y.;
RT "Identification of a member of mouse semaphorin family.";
RL FEBS Lett. 370:269-272(1995).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: EXPRESSED WIDELY IN THE NERVOUS TISSUES DURING
CC DEVELOPMENT. PREDOMINANTLY IN THE DEVELOPING BRAIN AND SPINAL
CC CORD.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: S79463; AAB35184.1; -
DR MGD: MGI:109252; SEMA4C.
DR INTERPRO: IPR001627; -
DR INTERPRO: IPR002165; -
DR INTERPRO: IPR003006; -
DR PFAM: PF01437; Plexin_repeat; 1.
DR PFAM: PF01403; Semaphorin; 1.
DR PFAM: PF00047; 19; 1.
DR Signal: Transmembrane; Immunoglobulin domain; Multigene family;
KW Neurogenesis; Developmental protein; Glycoprotein.
FT SIGNAL 1 21
FT CHAIN 22 834
FT DOMAIN 22 664
FT TRANSMEM 665 685
FT DOMAIN 686 834
FT DOMAIN 232 520
FT DOMAIN 571 635
FT DOMAIN 678 687
FT DOMAIN 709 774
FT DISULFID 578 628
FT CARBOHYD 106 106
FT CARBOHYD 121 121
FT CARBOHYD 310 310
FT CARBOHYD 419 419
FT CARBOHYD 522 522
FT CARBOHYD 565 565
FT SEQUENCE 834 AA; 92556 MW; 6868BB5F571482D CRC64;
SQ
Query Match 64.4%; Score 38; DB 1; Length 834;
Best Local Similarity 66.7%; Pred. No. 67;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 HASCPPGAG 9
Db 750 HARCQPGGG 758

RESULT 12
NTC4_MOUSE STANDARD; PRT; 1964 AA.
AC 31695; Q62389;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 4 PRECURSOR (TRANSFORMING

```

```
FT DOMAIN 153 EGF-LIKE 4. 189 EGF-LIKE 5. CALCIIUM-BINDING (POTENTIAL). 514 DISULFID 525 BY SIMILARITY.
FT DOMAIN 191 EGF-LIKE 5. 229 EGF-LIKE 6. 534 DISULFID 534 BY SIMILARITY.
FT DOMAIN 231 271 EGF-LIKE 7. 545 DISULFID 545 BY SIMILARITY.
FT DOMAIN 273 309 EGF-LIKE 8. 552 DISULFID 552 BY SIMILARITY.
FT DOMAIN 311 350 EGF-LIKE 9. 557 DISULFID 557 BY SIMILARITY.
FT DOMAIN 352 388 EGF-LIKE 10. 572 DISULFID 572 BY SIMILARITY.
FT DOMAIN 389 427 EGF-LIKE 11. 583 DISULFID 583 BY SIMILARITY.
FT DOMAIN 429 470 EGF-LIKE 12. 590 DISULFID 590 BY SIMILARITY.
FT DOMAIN 472 508 EGF-LIKE 13. 595 DISULFID 595 BY SIMILARITY.
FT DOMAIN 510 546 EGF-LIKE 14. 601 DISULFID 601 BY SIMILARITY.
FT DOMAIN 548 584 EGF-LIKE 15. 612 DISULFID 612 BY SIMILARITY.
FT DOMAIN 586 622 EGF-LIKE 16. 626 DISULFID 626 BY SIMILARITY.
FT DOMAIN 622 656 EGF-LIKE 17. 631 DISULFID 631 BY SIMILARITY.
FT DOMAIN 658 686 EGF-LIKE 18. 648 DISULFID 648 BY SIMILARITY.
FT DOMAIN 688 724 EGF-LIKE 19. 662 DISULFID 662 BY SIMILARITY.
FT DOMAIN 726 762 EGF-LIKE 20. 664 DISULFID 664 BY SIMILARITY.
FT DOMAIN 764 800 EGF-LIKE 21. 676 DISULFID 676 BY SIMILARITY.
FT DOMAIN 803 839 EGF-LIKE 22. 692 DISULFID 692 BY SIMILARITY.
FT DOMAIN 841 877 EGF-LIKE 23. 697 DISULFID 697 BY SIMILARITY.
FT DOMAIN 878 924 EGF-LIKE 24. 714 DISULFID 714 BY SIMILARITY.
FT DOMAIN 926 962 EGF-LIKE 25. 723 DISULFID 723 BY SIMILARITY.
FT DOMAIN 964 1000 EGF-LIKE 26. 735 DISULFID 735 BY SIMILARITY.
FT DOMAIN 1002 1040 EGF-LIKE 27. 741 DISULFID 741 BY SIMILARITY.
FT DOMAIN 1042 1081 EGF-LIKE 28. 752 DISULFID 752 BY SIMILARITY.
FT DOMAIN 1083 1122 EGF-LIKE 29. 761 DISULFID 761 BY SIMILARITY.
FT DOMAIN 1126 1167 3 X LIN/NOTCH REPEATS. 768 DISULFID 768 BY SIMILARITY.
FT REPEAT 1168 1208 LIN/NOTCH 1. 773 DISULFID 773 BY SIMILARITY.
FT REPEAT 1209 1242 LIN/NOTCH 2. 788 DISULFID 788 BY SIMILARITY.
FT REPEAT 1243 1282 LIN/NOTCH 3. 790 DISULFID 790 BY SIMILARITY.
FT DOMAIN 1572 1785 6 X ANK MOTIF REPEATS. 807 DISULFID 807 BY SIMILARITY.
FT REPEAT 1572 1603 ANK MOTIF 1. 812 DISULFID 812 BY SIMILARITY.
FT REPEAT 1603 1633 ANK MOTIF 2. 827 DISULFID 827 BY SIMILARITY.
FT REPEAT 1633 1653 ANK MOTIF 3. 838 DISULFID 838 BY SIMILARITY.
FT REPEAT 1653 1685 ANK MOTIF 4. 845 DISULFID 845 BY SIMILARITY.
FT REPEAT 1685 1719 ANK MOTIF 5. 856 DISULFID 856 BY SIMILARITY.
FT REPEAT 1719 1752 ANK MOTIF 6. 865 DISULFID 865 BY SIMILARITY.
FT REPEAT 1752 1785 ANK MOTIF 6. 876 DISULFID 876 BY SIMILARITY.
FT DISULFID 25 38 BY SIMILARITY. 882 DISULFID 882 BY SIMILARITY.
FT DISULFID 32 48 BY SIMILARITY. 897 DISULFID 897 BY SIMILARITY.
FT DISULFID 50 59 BY SIMILARITY. 903 DISULFID 903 BY SIMILARITY.
FT DISULFID 65 77 BY SIMILARITY. 912 DISULFID 912 BY SIMILARITY.
FT DISULFID 71 100 BY SIMILARITY. 923 DISULFID 923 BY SIMILARITY.
FT DISULFID 102 111 BY SIMILARITY. 930 DISULFID 930 BY SIMILARITY.
FT DISULFID 119 130 BY SIMILARITY. 935 DISULFID 935 BY SIMILARITY.
FT DISULFID 124 140 BY SIMILARITY. 952 DISULFID 952 BY SIMILARITY.
FT DISULFID 142 151 BY SIMILARITY. 968 DISULFID 968 BY SIMILARITY.
FT DISULFID 157 168 BY SIMILARITY. 979 DISULFID 979 BY SIMILARITY.
FT DISULFID 162 177 BY SIMILARITY. 988 DISULFID 988 BY SIMILARITY.
FT DISULFID 179 188 BY SIMILARITY. 990 DISULFID 990 BY SIMILARITY.
FT DISULFID 195 208 BY SIMILARITY. 999 DISULFID 999 BY SIMILARITY.
FT DISULFID 202 217 BY SIMILARITY. 1006 DISULFID 1006 BY SIMILARITY.
FT DISULFID 219 228 BY SIMILARITY. 1019 DISULFID 1019 BY SIMILARITY.
FT DISULFID 235 246 BY SIMILARITY. 1028 DISULFID 1028 BY SIMILARITY.
FT DISULFID 240 259 BY SIMILARITY. 1030 DISULFID 1030 BY SIMILARITY.
FT DISULFID 261 270 BY SIMILARITY. 1039 DISULFID 1039 BY SIMILARITY.
FT DISULFID 277 288 BY SIMILARITY. 1046 DISULFID 1046 BY SIMILARITY.
FT DISULFID 282 297 BY SIMILARITY. 1051 DISULFID 1051 BY SIMILARITY.
FT DISULFID 299 308 BY SIMILARITY. 1059 DISULFID 1059 BY SIMILARITY.
FT DISULFID 315 329 BY SIMILARITY. 1069 DISULFID 1069 BY SIMILARITY.
FT DISULFID 323 338 BY SIMILARITY. 1071 DISULFID 1071 BY SIMILARITY.
FT DISULFID 340 349 BY SIMILARITY. 1080 DISULFID 1080 BY SIMILARITY.
FT DISULFID 356 367 BY SIMILARITY. 1087 DISULFID 1087 BY SIMILARITY.
FT DISULFID 361 376 BY SIMILARITY. 1092 DISULFID 1092 BY SIMILARITY.
FT DISULFID 378 387 BY SIMILARITY. 1092 DISULFID 1092 BY SIMILARITY.
FT DISULFID 393 404 BY SIMILARITY. 1110 DISULFID 1110 BY SIMILARITY.
FT DISULFID 398 415 BY SIMILARITY. 1112 DISULFID 1112 BY SIMILARITY.
FT DISULFID 413 426 BY SIMILARITY.
FT DISULFID 437 449 BY SIMILARITY.
FT DISULFID 443 458 BY SIMILARITY.
FT DISULFID 460 469 BY SIMILARITY.
FT DISULFID 476 487 BY SIMILARITY.
FT DISULFID 481 496 BY SIMILARITY.
FT DISULFID 498 507 BY SIMILARITY.
```

Query Match 64.4% Score 38: DB 1: Length 1964;

Best Local Similarity 71.4% Pred No. 1.5e+02;

Matches 5: Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HASCPPG 7

Db 987 HCACPPG 993

RESULT 13

LMA5_MOUSE ID LMA5_MOUSE STANDARD; PRT; 3635 AA.

AC Q61001;

DT 01-NOV-1997 (Rel. 35, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE LAMININ ALPHA-5 CHAIN (FRAGMENT).

GN LMA5.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6 X CBA; TISSUE=LUNG;

RX MEDLINE=96081906; PubMed=7499364;
RA Miner J.H., Lewis R.M., Sanes J.R.;
RT "Molecular cloning of a novel laminin chain, alpha 5, and widespread
RL expression in adult mouse tissues.";
RN J. Biol. Chem. 270:28523-28526(1995).
RP (2)
REVIEWS.
RA Miner J.H., Lewis R.M., Sanes J.R.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
CC IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
CC -1- FUNCTION: ALPHA-5 CHAIN MAY BE THE MAJOR LAMININ ALPHA CHAIN OF
CC ADULT EPITHELIAL AND/OR ENDOTHELIAL BASAL LAMINAE.
CC -1- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT
CC MEMBRANES (MAJOR COMPONENT).
CC -1- TISSUE SPECIFICITY: IN ADULT, HIGH LEVELS IN HEART, LUNG, AND
CC KIDNEY; LOWER IN BRAIN, MUSCLE AND TESTIS; VERY LOW IN LIVER, GUT
CC AND SKIN. EXPRESSED IN MANY TISSUES IN EMBRYONIC DAY 11.
CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC -1- DOMAIN: DOMAINS VI, IV AND G ARE GLOBULAR.
CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC -1- SIMILARITY: CONTAINS 21.5 LAMININ EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 LAMININ DOMAINS IV.
CC -1- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.
CC -1- SIMILARITY: ALPHA-5 CHAIN IS MORE RELATED TO DROSOPHILA ALPHA
CC CHAIN THAN TO MAMMALIAN ALPHA CHAINS 1-4.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL: U37501; AAC53430.1; -;
CC HSP: P02468; ITLE.
DR MGD: MGI:105382; LAMA5.
DR INTERPRO: IPR000034; -;
DR INTERPRO: IPR000561; -;
DR INTERPRO: IPR001791; -;
DR INTERPRO: IPR001886; -;
DR INTERPRO: IPR002049; -;
DR PFAM: PF000052; laminin_B; 1.
DR PFAM: PF000053; laminin_EGF; 19.
DR PFAM: PF000054; laminin_G; 2.
DR PFAM: PF000055; laminin_Nterm; 1.
DR PROSITE: PS00022; EGF_1; 19.
DR PROSITE: PS01186; EGF_2; 3.
DR PROSITE: PS01248; LAMININ_TYPE_EGF; 19.
KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
Laminin EGF-like domain; Cell adhesion; Repeat.
FT NON_TER 1
FT DOMAIN <1 221 LAMININ N-TERMINAL (DOMAIN VI).
FT DOMAIN 222 772 10.5 X LAMININ EGF-LIKE REPEATS (DOMAIN
FT V).
FT DOMAIN 222 280 LAMININ EGF-LIKE 1.
FT DOMAIN 281 350 LAMININ EGF-LIKE 2.
FT DOMAIN 351 396 LAMININ EGF-LIKE 3.
FT DOMAIN 417 463 LAMININ EGF-LIKE 4.
FT DOMAIN 464 509 LAMININ EGF-LIKE 5.
FT DOMAIN 510 554 LAMININ EGF-LIKE 6.
FT DOMAIN 555 599 LAMININ EGF-LIKE 7.
FT DOMAIN 600 645 LAMININ EGF-LIKE 8.
FT DOMAIN 646 698 LAMININ EGF-LIKE 9.
FT DOMAIN 699 750 LAMININ EGF-LIKE 10.

FT DOMAIN 751 772 LAMININ EGF-LIKE 11 (INCOMPLETE).
FT DOMAIN 773 1359 LAMININ DOMAIN IV 1 (DOMAIN IV B).
FT DOMAIN 1360 1559 4.5 X LAMININ EGF-LIKE REPEATS (DOMAIN
FT III B).
FT DOMAIN 1360 1405 LAMININ EGF-LIKE 12.
FT DOMAIN 1406 1449 LAMININ EGF-LIKE 13.
FT DOMAIN 1450 1498 LAMININ EGF-LIKE 14.
FT DOMAIN 1499 1549 LAMININ EGF-LIKE 15.
FT DOMAIN 1550 1559 LAMININ EGF-LIKE 16 (N-TERMINAL).
FT DOMAIN 1560 1748 LAMININ DOMAIN IV 2 (DOMAIN IV A).
FT DOMAIN 1749 2085 6.5 X LAMININ EGF-LIKE REPEATS (DOMAIN
FT III A).
FT DOMAIN 1749 1781 LAMININ EGF-LIKE 16 (C-TERMINAL).
FT DOMAIN 1782 1831 LAMININ EGF-LIKE 17.
FT DOMAIN 1832 1887 LAMININ EGF-LIKE 18.
FT DOMAIN 1888 1941 LAMININ EGF-LIKE 19.
FT DOMAIN 1942 1988 LAMININ EGF-LIKE 20.
FT DOMAIN 1989 2035 LAMININ EGF-LIKE 21.
FT DOMAIN 2036 2085 LAMININ EGF-LIKE 22.
FT DOMAIN 2086 2678 LAMININ II AND I.
FT DOMAIN 2679 3635 5 X LAMININ G-LIKE REPEATS (DOMAIN G).
FT DOMAIN 2679 2849 LAMININ G-LIKE 1.
FT DOMAIN 2850 3035 LAMININ G-LIKE 2.
FT DOMAIN 3036 3241 LAMININ G-LIKE 3.
FT DOMAIN 3242 3427 LAMININ G-LIKE 4.
FT DOMAIN 3428 3635 LAMININ G-LIKE 5.
FT DOMAIN 2122 2174 COILED COIL (POTENTIAL).
FT DOMAIN 2247 2381 COILED COIL (POTENTIAL).
FT DOMAIN 2521 2538 COILED COIL (POTENTIAL).
FT DOMAIN 2556 2622 COILED COIL (POTENTIAL).
FT SITE 1640 CELL ATTACHMENT SITE (POTENTIAL).
FT SITE 1756 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 222 231 BY SIMILARITY.
FT DISULFID 224 244 BY SIMILARITY.
FT DISULFID 246 255 BY SIMILARITY.
FT DISULFID 258 278 BY SIMILARITY.
FT DISULFID 281 290 BY SIMILARITY.
FT DISULFID 318 327 BY SIMILARITY.
FT DISULFID 330 348 BY SIMILARITY.
FT DISULFID 351 362 BY SIMILARITY.
FT DISULFID 353 369 BY SIMILARITY.
FT DISULFID 371 380 BY SIMILARITY.
FT DISULFID 383 393 BY SIMILARITY.
FT DISULFID 417 429 BY SIMILARITY.
FT DISULFID 419 438 BY SIMILARITY.
FT DISULFID 440 449 BY SIMILARITY.
FT DISULFID 452 461 BY SIMILARITY.
FT DISULFID 464 476 BY SIMILARITY.
FT DISULFID 466 483 BY SIMILARITY.
FT DISULFID 485 494 BY SIMILARITY.
FT DISULFID 497 507 BY SIMILARITY.
FT DISULFID 510 522 BY SIMILARITY.
FT DISULFID 512 528 BY SIMILARITY.
FT DISULFID 530 539 BY SIMILARITY.
FT DISULFID 542 552 BY SIMILARITY.
FT DISULFID 555 567 BY SIMILARITY.
FT DISULFID 557 573 BY SIMILARITY.
FT DISULFID 575 584 BY SIMILARITY.
FT DISULFID 587 597 BY SIMILARITY.
FT DISULFID 600 612 BY SIMILARITY.
FT DISULFID 602 619 BY SIMILARITY.
FT DISULFID 621 630 BY SIMILARITY.
FT DISULFID 633 643 BY SIMILARITY.
FT DISULFID 1360 1372 BY SIMILARITY.
FT DISULFID 1362 1379 BY SIMILARITY.
FT DISULFID 1381 1390 BY SIMILARITY.
FT DISULFID 1393 1403 BY SIMILARITY.
FT DISULFID 1450 1465 BY SIMILARITY.
FT DISULFID 1452 1472 BY SIMILARITY.
FT DISULFID 1474 1483 BY SIMILARITY.
FT DISULFID 1486 1496 BY SIMILARITY.
FT DISULFID 1499 1511 BY SIMILARITY.

Search completed: January 30, 2001, 17:06:14
Job time: 1235 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:55:02 ; Search time 254.27 Seconds
(without alignments)
4.610 Million cell updates/sec

Title: US-09-518-931-4_COPY_129_138

Perfect score: 59

Sequence: 1 HASCPPGAGV 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_15:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	59	100.0	300	4	Q95407
2	51	86.4	401	11	O08712
3	49	83.1	401	11	O08727
4	48	81.4	372	4	Q9UHP4
5	48	81.4	401	4	O00300
6	43	72.9	234	4	Q95851
7	43	72.9	241	4	Q9Y505
8	43	72.9	255	4	Q9NVJ9
9	42	71.2	47	5	Q9VBF3
10	41	69.5	299	2	Q9RVG0
11	40	67.8	278	1	O28227
12	40	67.8	302	13	Q9PUS0
13	40	67.8	1337	2	Q9KJN7
14	40	67.8	2482	6	Q28833
15	40	67.8	2813	6	Q28311
16	40	67.8	2813	6	Q9TS14
17	39.5	66.9	453	4	O75040
18	39	66.1	177	2	Q9R8X2
19	39	66.1	177	2	Q9R8X1

20	39	66.1	177	2	Q9R8X0
21	39	66.1	177	2	Q9R8W8
22	39	66.1	178	2	Q9Z3N4
23	39	66.1	179	11	Q9JKW9
24	39	66.1	223	4	Q9UBS9
25	39	66.1	383	5	O62543
26	39	66.1	395	2	Q9RYI2
27	39	66.1	752	12	Q9QLK0
28	39	66.1	2261	4	Q9NOV4
29	38	64.4	281	2	Q9K413
30	38	64.4	480	2	O53244
31	38	64.4	510	4	Q9NX92
32	38	64.4	512	10	Q9MA38
33	38	64.4	594	5	Q9VEP4
34	38	64.4	686	13	Q9IA83
35	38	64.4	686	13	Q9IA82
36	38	64.4	981	4	Q9NTI0
37	38	64.4	1095	4	Q99458
38	38	64.4	1323	5	Q9NJ14
39	38	64.4	1343	5	Q9NJ16
40	38	64.4	1436	5	Q9VH03
41	38	64.4	1534	5	Q9WIA8
42	38	64.4	1631	4	Q9Y606
43	38	64.4	1696	5	Q9NJ15
44	38	64.4	1964	11	O35442
45	38	64.4	1999	4	Q99940

ALIGNMENTS

RESULT 1

Q95407	PRELIMINARY;	PRT;	300 AA.
ID	Q95407		
AC	Q95407;		
DT	01-MAY-1999 (TReMBLrel. 10, Created)		
DT	01-MAY-1999 (TReMBLrel. 10, Last sequence update)		
DE	01-OCT-2000 (TReMBLrel. 15, Last annotation update)		
DE	DECOY RECEPTOR 3 (M68) (M68C) (M68E).		
GN	DCR3 OR TR6.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=99087326; PubMed=9872321;		
RA	Pitti R.N., Marsters S.A., Lawrence D.A., Roy M., Kischkel F.C.,		
RA	Dowd P., Huang A., Donahue C.J., Sherwood S.W., Baldwin D.T.,		
RA	Godowski P.J., Wood W.I., Gurney A.L., Hillan K.J., Cohen R.L.,		
RA	Goddard A.D., Botstein D., Ashkenazi A.;		
RT	"Genomic amplification of a decoy receptor for Fas ligand in lung and		
RT	colon cancer.";		
RL	Nature 396:699-703(1998).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=BLOOD;		
RX	MEDLINE=99253915; PubMed=10318773;		
RA	Yu K.Y., Kwon B., Ni J., Zhai Y., Ebner R., Kwon B.S.;		
RT	"A newly identified member of tumor necrosis factor receptor		
RT	superfamily (TR6) suppresses LIGHT-mediated apoptosis.";		
J.	Biol. Chem. 274:13733-13736(1999).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=PANCREAS;		
RX	MEDLINE=20122600; PubMed=10655513;		
RA	Bai C., Connolly B., Metzker M.L., Hilliard C.A., Liu X., Sandig V.,		
RA	Soderman A., Galloway S.M., Liu O., Austin C.P., Caskey C.T.;		
RT	"Overexpression of M68/Dcr3 in human gastrointestinal tract tumors		
RT	independent of gene amplification and its location in a four-gene		
RT	cluster.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 97:1230-1235(2000).		
DR	EMBL; AF104419; AAD03056.1; -.		

Q9r8x0	rhizobium s
Q9r8w8	rhizobium s
Q9z3n4	rhizobium s
Q9jkw9	mus musculus
Q9uds9	homo sapien
O62543	geodia cydo
Q9y12	deinococcus
Q9q1k0	influenza b
Q9nqv4	homo sapien
Q9k413	streptomyce
O53244	mycobacteri
Q9nxx2	homo sapien
Q9ma38	arabidopsis
Q9vep4	drosophila
Q9ia83	paralichthy
Q9ia82	paralichthy
Q9nti0	homo sapien
Q99458	homo sapien
Q9nj14	branchiosto
Q9nj16	branchiosto
Q9vh03	drosophila
Q9w1a8	drosophila
Q9y606	homo sapien
Q9nj15	branchiosto
O35442	mus musculo
Q99940	homo sapien

DR	EMBL; AF134240; AAD29688.1; -	CC	-1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR	EMBL; AF217796; AAF35244.1; -	CC	EMBL; U94331; AAB53708.1; -
DR	EMBL; AF217793; AAF33685.1; -	DR	EMBL; AB013898; BAA28269.1; -
DR	EMBL; AF217794; AAF33686.1; -	DR	EMBL; AB013903; BAA33388.1; -
DR	HSSP; P25942; ICDF	DR	EMBL; AB013899; BAA33388.1; JOINED.
DR	INTERPRO; IPR000561; -	DR	EMBL; AB013900; BAA33388.1; JOINED.
DR	INTERPRO; IPR001368; -	DR	EMBL; AB013901; BAA33388.1; JOINED.
DR	PFAM; PF00020; TNFR_C6; 4.	DR	EMBL; AB013902; BAA33388.1; JOINED.
DR	PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.	DR	HSSP; P25942; ICDF.
DR	PROSITE; PS01186; EGF_2; UNKNOWN_1.	DR	MGI; MGI:109587; Opg.
DR	PROSITE; PS50050; TNFR_NGFR_2; 2.	DR	INTERPRO; IPR000488; -
DR	PRODOM; PD000771; -; 1.	DR	INTERPRO; IPR001368; -
DR	Receptor.	DR	PFAM; PF00020; TNFR_C6; 3.
DR	SEQUENCE 300 AA; 32679 MW; F90AE33718449AF CRC64;	DR	PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR		DR	PROSITE; PS50050; TNFR_NGFR_2; 2.
DR		DR	PRODOM; PD000771; -; 1.
DR		DR	Glycoprotein; Repeat; Cytokine; Signal.
DR		DR	FT SIGNAL 1 21
DR		DR	FT CHAIN 22 401
DR		DR	FT DOMAIN 23 201
DR		DR	FT REPEAT 23 63
DR		DR	FT REPEAT 64 106
DR		DR	FT REPEAT 107 143
DR		DR	FT REPEAT 144 201
DR		DR	FT DOMAIN 306 365
DR		DR	FT DISULFID 41 54
DR		DR	FT DISULFID 44 62
DR		DR	FT DISULFID 65 80
DR		DR	FT DISULFID 83 97
DR		DR	FT DISULFID 87 105
DR		DR	FT DISULFID 118 142
DR		DR	FT DISULFID 145 160
DR		DR	FT CARBOHYD 98 98
DR		DR	FT CARBOHYD 165 165
DR		DR	FT CARBOHYD 178 178
DR		DR	FT CARBOHYD 289 289
DR		DR	FT VARIANT 138 138
DR		DR	FT VARIANT 161 161
DR		DR	FT VARIANT 165 165
DR		DR	FT VARIANT 288 288
DR		DR	FT VARIANT 296 296
DR		DR	FT SEQUENCE 401 AA; 45923 MW; CAA6102D3B312470 CRC64;
DR		DR	SEQUENCE 401 AA; 45923 MW; CAA6102D3B312470 CRC64;
DR		DR	Query Match 86.4%; Score 51; DB 11; Length 401;
DR		DR	Best Local Similarity 80.0%; Pred. No. 0.73;
DR		DR	Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
DR		DR	QY 1 HASCPPGAGV 10
DR		DR	DB 121 HRCPPGSGV 130
DR		DR	RESULT 3
DR		DR	O08727 PRELIMINARY; PRT; 401 AA.
DR		DR	AC O08727; (TREMblrel. 04, Created)
DR		DR	DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)
DR		DR	DT 01-JUL-1997 (TREMblrel. 15, Last annotation update)
DR		DR	DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
DR		DR	DE OSTROPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)
DR		DR	DE (OCIF).
DR		DR	GN TNFRSF11B OR OPG.
DR		DR	OS Rattus norvegicus (Rat).
DR		DR	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DR		DR	OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

```

OX NCBI_TaxID=10116;
RN {1}
RP SEQUENCE FROM N.A.
RC TISSUE=INTESTINE;
RX MEDLINE=97262071; PubMed=9108485;
RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,
Luethy R., Nguyen H.Q., Woodson S., Bennett L., Boone T., Shimamoto G.,
Derose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,
Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,
Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,
Suggs S., Boyle W.J.;
RA "Osteoprotegerin: a novel secreted protein involved in the regulation
of bone density.";
RL Cell 89:309-319(1997).
CC -1- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES
OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY
SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO
OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN
STROMAL CELLS AND OSTEOCLAST PROGENITORS.
CC -1- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM) (BY
SIMILARITY).
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL; U94330; AAB53707.1; -.
DR HSP; P25942; ICDF.
DR INTERPRO: IPR001368; -.
DR PFAM: PF00020; TNFR_c6; 4.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PRODOM; PD000771; -.
KW Glycoprotein; Repeat; Cytokine; Signal.
FT SIGNAL 1 21 BY SIMILARITY.
FT CHAIN 22 401 OSTEOPROTEGERIN.
FT DOMAIN 23 201 4 X TNFR-CYS.
FT REPEAT 23 63 TNFR-CYS 1.
FT REPEAT 64 106 TNFR-CYS 2.
FT REPEAT 107 143 TNFR-CYS 3.
FT REPEAT 144 201 TNFR-CYS 4.
FT DOMAIN 306 365 DEATH DOMAIN.
FT DISULFID 41 54 BY SIMILARITY.
FT DISULFID 44 62 BY SIMILARITY.
FT DISULFID 65 80 BY SIMILARITY.
FT DISULFID 83 97 BY SIMILARITY.
FT DISULFID 87 105 BY SIMILARITY.
FT DISULFID 118 142 BY SIMILARITY.
FT DISULFID 145 160 BY SIMILARITY.
FT CARBOHYD 98 98 POTENTIAL.
FT CARBOHYD 165 165 POTENTIAL.
FT CARBOHYD 178 178 POTENTIAL.
FT CARBOHYD 289 289 POTENTIAL.
SQ SEQUENCE 401 AA; 46192 MW; FEC6A31F1D4E573A CRC64;

Query Match 83.1%; Score 49; DB 11; Length 401;
Best Local Similarity 80.0%; Pred. No. 1.5;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 HASCPPPGAGV 10
Db 121 HRSPPGLGV 130
1 ||||| 11

RESULT 4
Q9UHP4 PRELIMINARY; PRT; 372 AA.
AC Q9UHP4
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE OSTEOPROTEGERIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
CC -1- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES

```

```

OX NCBI_TaxID=9606;
RN {1}
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX He Z.-Y., Yang G.-Z., Zhang W.-J., Wu X.-F.;
RT "Cloning and Expression of Osteoprotegerin from Homo sapiens.";
RL Sheng Wu Hua Hsueh Yu Sheng Wu Li Hsueh Pao 31:680-684(1999).
RA EMBL; AF134187; AAF20168.1; -.
DR HSP; P25942; ICDF.
DR INTERPRO: IPR001368; -.
DR PFAM; PF00020; TNFR_c6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
DR PROSITE; PS00500; TNFR_NGFR_2; 2.
FT NON_TER 1 1
SQ SEQUENCE 372 AA; 42758 MW; F02527A5CD01CCD3 CRC64;

Query Match 81.4%; Score 48; DB 4; Length 372;
Best Local Similarity 80.0%; Pred. No. 2.1;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 HASCPPPGAGV 10
Db 100 HRSPPGFGV 109
1 ||||| 11

RESULT 5
O00300 PRELIMINARY; PRT; 401 AA.
AC O00300; O60236;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE OSTEOPROTEGERIN PRECURSOR (OSTEOCLASTOGENESIS INHIBITORY FACTOR)
DE (OCIF) (TUMOR NECROSIS FACTOR RECEPTOR SUPERFAMILY MEMBER IIB).
GN TNFRSF11B OR OPG OR OCIF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN {1}
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RX MEDLINE=97262071; PubMed=9108485;
RA Simonet W.S., Lacey D.L., Dunstan C.R., Kelley M., Chang M.-S.,
Luethy R., Nguyen H.Q., Woodson S., Bennett L., Boone T., Shimamoto G.,
Derose M., Elliott R., Colombero A., Tan H.-L., Trail G., Sullivan J.,
Davy E., Bucay N., Renshaw-Gegg L., Hughes T.M., Hill D., Pattison W.,
Campbell P., Sander S., Van G., Tarpley J., Derby P., Lee R.,
Suggs S., Boyle W.J.;
RA "Osteoprotegerin: a novel secreted protein involved in the regulation
of bone density.";
RL Cell 89:309-319(1997).
RN {2}
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG FIBROBLAST;
RX MEDLINE=98151033; PubMed=9492069;
RA Yasuda H., Shima N., Nakagawa N., Mochizuki S.-I., Yano K., Fujise N.,
Sato Y., Goto M., Yamaguchi K., Kuriyama M., Kanno T., Murakami A.,
Tsuda E., Morinaga T., Higashio K.;
RT "Identity of osteoclastogenesis inhibitory factor (OCIF) and
osteoprotegerin (OPG): a mechanism by which OPG/OCIF inhibits
osteoclastogenesis in vitro.";
RL Endocrinology 139:1329-1337(1998).
RN {3}
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=98351569; PubMed=9688283;
RA Morinaga T., Nakagawa N., Yasuda H., Tsuda E., Higashio K.;
RT "Cloning and characterization of the gene encoding human
osteoprotegerin/osteoclastogenesis-inhibitory factor.";
RL Eur. J. Biochem. 254:685-691(1998).
CC -1- FUNCTION: MAY REGULATE BONE RESORPTION BY BLOCKING TERMINAL STAGES

```

CC OF OSTEOCLAST DIFFERENTIATION AND ALLOWING ACCUMULATION OF NEWLY
CC SYNTHESIZED BONE AND CARTILAGE. MAY INHIBIT IN VITRO
CC OSTEOCLASTOGENESIS BY INTERRUPTING CELL-TO-CELL SIGNALING BETWEEN
CC STROMAL CELLS AND OSTEOCLAST PROGENITORS.
CC -|- SUBUNIT: HOMODIMER (MAJOR FORM) AND MONOMER (MINOR FORM) (BY
CC SIMILARITY).
CC -|- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -|- TISSUE SPECIFICITY: EXPRESSED AT HIGHEST LEVELS IN LUNG, HEART,
CC KIDNEY, PLACENTA, THYROID, SPINAL CORD AND LIVER. ALSO DETECTED IN
CC A NUMBER OF OTHER HEMATOPOIETIC AND IMMUNE ORGANS. NOT DETECTED IN
CC THE PANCREAS OR PERIPHERAL BLOOD LYMPHOCYTES.
CC -|- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC EMBL: AB002146; BAA25910.1; -
CC EMBL: AB008822; BAA32076.1; -
CC EMBL: AB008821; BAA32076.1; JOINED.
CC EMBL: U94332; AAB53709.1; -
CC HSP: P25942; 1CDF.
CC MIM: 602643; -
CC INTERPRO: IPR001368; -
CC PFM: PF00020; TNFR_c6; 3.
CC PROSITE: PS00652; TNFR_NGFR_1; UNKNOWN_1.
CC PROSITE: PS00050; TNFR_NGFR_2; 2.
CC PRODOM: PD000771; -; 1.
CC Glycoprotein; Repeat; Cytokine; Signal.
CC SIGNAL 1 21 BY SIMILARITY.
CC CHAIN 22 401 OSTEOPROTEGERIN.
CC DOMAIN 23 183 4 X TNFR-CYS.
CC REPEAT 23 63 TNFR-CYS 1.
CC REPEAT 64 106 TNFR-CYS 2.
CC REPEAT 107 143 TNFR-CYS 3.
CC REPEAT 144 201 TNFR-CYS 4.
CC DOMAIN 306 365 DEATH DOMAIN.
CC DISULFID 41 54 BY SIMILARITY.
CC DISULFID 44 62 BY SIMILARITY.
CC DISULFID 65 80 BY SIMILARITY.
CC DISULFID 83 97 BY SIMILARITY.
CC DISULFID 87 105 BY SIMILARITY.
CC DISULFID 118 142 BY SIMILARITY.
CC DISULFID 145 160 BY SIMILARITY.
CC CARBOHYD 98 28 POTENTIAL.
CC CARBOHYD 152 152 POTENTIAL.
CC CARBOHYD 165 165 POTENTIAL.
CC CARBOHYD 178 178 POTENTIAL.
CC CARBOHYD 289 289 POTENTIAL.
CC CONFLICT 263 263 A -> D (IN REF. 2 AND 3).
CC SEQUENCE 401 AA; 45996 MW; EB42FA51C9D7C71E CRC64;

Query Match 81.4%; Score 48; DB 4; Length 401;
Best Local Similarity 80.0%; Pred. No. 2.2;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HASCPPGAGV 10
| | | | | | | |
DB 121 HRSCPPGFGV 130

RESULT 6
O95851 ID O95851 PRELIMINARY; PRT; 234 AA.
AC O95851;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE TNF RECEPTOR SUPERFAMILY ACTIVATION-INDUCIBLE PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99156876; PubMed=10037686;
RA Kwon B., Yu K.Y., Ni J., Yu G.L., Jang I.K., Kim Y.J., Xing L.,

RA Liu D., Wang S.X., Kwon B.S.;
RT "Identification of a novel activation-inducible protein of the tumor
RT necrosis factor receptor superfamily and its ligand.";
RL J. Biol. Chem. 274:6056-6061(1999).
KW EMBL: AF117297; AAD19694.1; -
KW Receptor.
SQ SEQUENCE 234 AA; 25110 MW; 55DB593CD0B1DB45 CRC64;

Query Match 72.9%; Score 43; DB 4; Length 234;
Best Local Similarity 70.0%; Pred. No. 9.1;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 HASCPPGAGV 10
| | | | | | | |
DB 91 HHPCPPGGGV 100

RESULT 7
O9Y5U5 ID O9Y5U5 PRELIMINARY; PRT; 241 AA.
AC O9Y5U5;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE GLUCOCORTICOID-INDUCED TNFR-RELATED PROTEIN.
GN TNFRSF18.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BONE MARROW;
RA Gurney A.L., Marsters S.A., Huang A., Pitti R.M., Mark M.,
RA Baldwin D.T., Gray A.M., Dowd P., Brush J., Heldens S., Schow P.,
RA Goddard A.D., Wood W.I., Baker K.P., Godowski P.J., Ashkenazi A.;
RT "Identification of a new member of the tumor necrosis factor family
RT and its receptor, a human ortholog of mouse GITR.";
RL Curr. Biol. 0:0-0(1999).
DR EMBL: AF125304; AAD22635.1; -
SQ SEQUENCE 241 AA; 26000 MW; 90DC3B4AA7E82CBE CRC64;

Query Match 72.9%; Score 43; DB 4; Length 241;
Best Local Similarity 70.0%; Pred. No. 9.4;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 HASCPPGAGV 10
| | | | | | | |
DB 91 HHPCPPGGGV 100

RESULT 8
Q9NYJ9 ID Q9NYJ9 PRELIMINARY; PRT; 255 AA.
AC Q9NYJ9;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE GITR-D.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=THYMUS;
RA Nocentini G., Bartoli A., Ronchetti S., Mastroiaca E., Ricciardi C.;
RT "Identification of a soluble human GITR splicing (HGTR-D).";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF241229; AAF63506.1; -
SQ SEQUENCE 255 AA; 26827 MW; C986652AC97AF2CC CRC64;

Query Match 72.9%; Score 43; DB 4; Length 255;
 Best Local Similarity 70.0%; Pred. No. 9.8;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 HASCPPGAGV 10
 | | | | |
 Db 91 HHPCPPGQV 100

RESULT 9
 Q9VBF3 PRELIMINARY; PRT; 47 AA.
 AC Q9VBF3
 DT 01-MAY-2000 (TEMBLrel. 13, Created)
 DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TEMBLrel. 13, Last annotation update)
 DE CG6503 PROTEIN.
 GN CG6503.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abrial J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo S., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003756; AAF56589.1; .
 DR FLYBASE; FBgn040606; CG6503.
 SQ SEQUENCE 47 AA; 4829 MW; 7972B89E160EC591 CRC64;

Query Match 71.2%; Score 42; DB 5; Length 47;

Best Local Similarity 75.0%; Pred. No. 3.3;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 HASCPPGA 8
 | | | | |
 Db 28 HGSCPPGS 35

RESULT 10
 Q9RYG0 PRELIMINARY; PRT; 299 AA.
 AC Q9RYG0;
 DT 01-MAY-2000 (TEMBLrel. 13, Created)
 DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TEMBLrel. 14, Last annotation update)
 DE HYPOTHETICAL 31.7 KDA PROTEIN.
 GN DRA0358.
 OS Deinococcus radiodurans.
 OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
 OX NCBI_TaxID=1299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=R1;
 RX MEDLINE=20036896; PubMed=10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT "Genome Sequence of the Radioresistant Bacterium Deinococcus
 radiodurans R1";
 RL Science 286:1571-1577(1999).
 DR EMBL: AE001863; AAF12505.1; .
 DR TIGR: DRA0358; .
 KW Hypothetical protein.
 SQ SEQUENCE 299 AA; 31674 MW; B090C9B81907D033 CRC64;

Query Match 69.5%; Score 41; DB 2; Length 299;
 Best Local Similarity 72.7%; Pred. No. 24;
 Matches 8; Conservative 1; Mismatches 0; Indels 2; Gaps 1;

QY 1 HASC--PPGAG 9
 | | | | |
 Db 97 HAACTLPPGAG 107

RESULT 11
 O28227 PRELIMINARY; PRT; 278 AA.
 AC O28227;
 DT 01-JAN-1998 (TEMBLrel. 05, Created)
 DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
 DT 01-AUG-1998 (TEMBLrel. 07, Last annotation update)
 DE 2-KETOISOLEUATE FERREDOXIN OXIDOREDUCTASE, SUBUNIT BETA (VORB).
 GN AF2052.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
 OC Archaeoglobus.
 OX NCBI_TaxID=2234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=938475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., Hickey E.K., Peterson J.D.,
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Kyrpides N.C.,
 RA Richardson D.L., Kerlavage A.R., Graham G., Sutton G.G., Gill S.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Adams M.D., Loftus B.,
 RA Kirkness E.F., Dougherty B.A., McKenney K., Badger J.H., Zhou L.,
 RA Peterson S., Reich C.I., McNeil L.K., Weidman J.F., McDonald L., Utterback T.,
 RA Overbeek R., Gocayne J.D.,

RA Cotton M.D., Spriggs T., Artlach P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon *Archaeoglobus fulgidus*.";
 RL Nature 390:364-370(1997).
 DR EMBL: AE000961; AAB89197.1; -.
 DR TIGR: AF2052; -.
 KW Hypothetical protein.
 SQ SEQUENCE 278 AA; 30597 MW; A5DA2E0B8AB21D5 CRC64;

Query Match 67.8%; Score 40; DB 1; Length 278;
 Best Local Similarity 85.7%; Pred. No. 32;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HASCPPG 7
 Db 189 HAPCPPG 195

RESULT 12
 Q9PUS0 ID Q9PUS0 PRELIMINARY; PRT; 302 AA.
 AC Q9PUS0;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE DECOY RECEPTOR.
 OS Salvelinus fontinalis (Brook trout).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
 OX NCBI_TaxID=8038;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bobe J., Goetz F.W.;
 RT "A tumor necrosis factor receptor homolog is up-regulated in the brook
 RT trout (*Salvelinus fontinalis*) ovary at the completion of ovulation.";
 RL Biol. Reprod. 0:0-0(1999).
 DR EMBL: AF156738; AAD56428.1; -.
 DR HSSP: P19438; IEXT.
 DR INTERPRO: IPR000561; -.
 DR INTERPRO: IPR001368; -.
 DR PFAM: PF00020; TNFR_c6; 4.
 DR PROSITE: PS00652; TNFR_NGFR_1; UNKNOWN_1.
 DR PROSITE: PS0186; EGF_2; UNKNOWN_1.
 DR PROSITE: PS00050; TNFR_NGFR_2; 1.
 KW Receptor.
 SQ SEQUENCE 302 AA; 34037 MW; E44C73477F05C3DF CRC64;

Query Match 67.8%; Score 40; DB 13; Length 302;
 Best Local Similarity 66.7%; Pred. No. 35;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 HASCPPGAG 9
 Db 117 HKECPPGYG 125

RESULT 13
 Q9KJN7 ID Q9KJN7 PRELIMINARY; PRT; 1337 AA.
 AC Q9KJN7;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE PUTATIVE SERINE/THREONINE PROTEIN KINASE.
 OS PKND2.
 GN Myxococcus xanthus.
 OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;

OC Myxococcales; Cystobacterineae; Myxococcaceae; Myxococcus.
 OX NCBI_TaxID=34;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=D22;
 RX MEDLINE=20032045; PubMed=10564511;
 RA Cho K., Zusman D.R.;
 RT "Sporulation timing in *Myxococcus xanthus* is controlled by the *espAB*
 RT locus.";
 RL Mol. Microbiol. 34:714-725(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=D22;
 RA Cho K., Zusman D.R.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF163841; AAF87930.1; -.
 KW Kinase; Serine/threonine-protein kinase.
 SQ SEQUENCE 1337 AA; 144737 MW; 60B6E8CA4C92499D CRC64;

Query Match 67.8%; Score 40; DB 2; Length 1337;
 Best Local Similarity 85.7%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HASCPPG 7
 Db 15 HAPCPPG 21

RESULT 14
 Q28B33 ID Q28B33 PRELIMINARY; PRT; 2482 AA.
 AC Q28B33;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE VON WILLEBRAND FACTOR (FRAGMENT).
 GN VWF.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96376329; PubMed=8782436;
 RA Sjöberg A., Seaman W.T., Bellingier D.A., Griggs T.R., Nichols T.C.,
 RA Chowdhary B.P.;
 RT "FISH mapping of the porcine vWF gene to chromosome 5q21 extends
 RT syntenic homology with human chromosome 12.";
 RL Hereditas 124:199-202(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Seaman W.T., Read M.S., Bellingier D.A., Nichols T.C.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RX SEQUENCE OF 397-553 FROM N.A.
 RX MEDLINE=93356762; PubMed=8352759;
 RA Laverne J.M., Piao Y.C., Ferreira V., Kerbirou-Nabias D.,
 RA Bahnak B.R., Meyer D.;
 RT "Primary structure of the factor VIII binding domain of human, porcine
 RT and rabbit von Willebrand factor.";
 RL Biochem. Biophys. Res. Commun. 194:1019-1024(1993).
 DR EMBL: AF052036; AAC06229.1; -.
 DR HSSP: S64541; AAE27829.2; -.
 DR HSSP: F04275; IATZ.
 DR INTERPRO: IPR00359; -.
 DR INTERPRO: IPR000864; -.
 DR INTERPRO: IPR001007; -.
 DR INTERPRO: IPR001211; -.
 DR INTERPRO: IPR001846; -.
 DR INTERPRO: IPR001928; -.
 DR INTERPRO: IPR002035; -.
 DR INTERPRO: IPR002919; -.

DR PFAM; PF00007; Cys_knot; 1.
DR PFAM; PF00092; vwa; 3.
DR PFAM; PF00093; vwc; 3.
DR PFAM; PF00094; vwd; 3.
DR PFAM; PF01826; TIL; 3.
DR PRINTS; PR00292; POTATOINHTR.
DR PRINTS; PR00365; ENDOTHELIN.
DR PRINTS; PR00453; VWFADOMAIN.
DR PROSITE; PS00119; PA2_ASP; UNKNOWN_1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01208; VWFC; UNKNOWN_3.
DR PROSITE; PS01225; CTCK_2; 1.
FT NON_TER 1
SQ SEQUENCE 2482 AA; 272394 MW; D499B7DDFBBCAEDD CRC64;

Query Match 67.8%; Score 40; DB 6; Length 2482;
Best Local Similarity 85.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HASCPPG 7
|||
Db 848 HAHCPPG 854

RESULT 15
Q28311
ID Q28311 PRELIMINARY; PRT; 2813 AA.
AC Q28311;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE VON WILLEBRAND FACTOR.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Montgomery R.R.; Fahs S.; Montgomery M.W.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1234-1669 FROM N.A.
RC TISSUE=BLOOD;
RA Mancuso D.J.; Christopherson P.A.; Kroner P.A.; Montgomery R.R.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U66246; AAB93766.1; -.
DR EMBL; L16903; AAA30903.1; -.
DR HSSP; P04275; 1AUQ.
DR INTERPRO; IPR000004; -.
DR INTERPRO; IPR000359; -.
DR INTERPRO; IPR001007; -.
DR INTERPRO; IPR001211; -.
DR INTERPRO; IPR001846; -.
DR INTERPRO; IPR002035; -.
DR INTERPRO; IPR002919; -.
DR PFAM; PF00007; Cys_knot; 1.
DR PFAM; PF00092; vwa; 3.
DR PFAM; PF00093; vwc; 3.
DR PFAM; PF00094; vwd; 4.
DR PFAM; PF01826; TIL; 4.
DR PRINTS; PR00453; VWFADOMAIN.
DR PROSITE; PS00119; PA2_ASP; UNKNOWN_1.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01208; VWFC; 3.
DR PROSITE; PS01225; CTCK_2; 1.
SQ SEQUENCE 2813 AA; 309731 MW; 21FF4DDF46257AAA CRC64;

Query Match 67.8%; Score 40; DB 6; Length 2813;
Best Local Similarity 85.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HASCPPG 7
|||
Db 1174 HAHCPPG 1180

Search completed: January 30, 2001, 16:55:04
Job time: 920 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:47:51 ; Search time 132.77 Seconds
(without alignments)
1.352 Million cell updates/sec

Title: US-09-518-931-4_COPY_129_138

Perfect score: 59

Sequence: 1 HASCPPGAGV 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/PCRUS_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	59	100.0	300	2	US-08-794-796-2
2	51	86.4	401	3	US-08-974-022-4
3	49	83.1	401	3	US-08-974-022-2
4	48	81.4	401	3	US-08-974-022-6
5	46	78.0	346	2	US-08-602-359A-34
6	43	72.9	228	3	US-08-911-423-6
7	43	72.9	232	3	US-08-911-423-7
8	43	72.9	241	3	US-08-911-423-4
9	43	72.9	311	3	US-08-911-423-8
10	40	67.8	41	1	US-08-050-319B-39
11	40	67.8	41	1	US-08-465-982-39
12	40	67.8	162	2	US-08-219-237B-7
13	40	67.8	2050	2	US-08-347-594A-2
14	40	67.8	2813	3	US-08-896-449A-2
15	40	67.8	2813	3	US-09-132-652-2
16	39	66.1	1375	3	US-08-665-259-26
17	39	66.1	1375	3	US-08-762-500-26
18	37	62.7	427	3	US-08-448-722A-4
19	36	61.0	452	3	US-09-059-522-3
20	36	61.0	452	3	US-09-382-027-3
21	36	61.0	621	3	US-09-059-522-1
22	36	61.0	621	3	US-09-382-027-1
23	35	59.3	13	1	US-08-444-231-20
24	35	59.3	13	1	US-08-152-443A-20
25	35	59.3	119	2	US-08-219-237B-3
26	35	59.3	149	4	PCT-US95-17083-6
27	35	59.3	189	3	US-08-646-273-14
28	35	59.3	219	3	US-08-974-022-45

29	35	59.3	229	1	US-08-383-750-6
30	35	59.3	229	2	US-08-684-687-2
31	35	59.3	229	3	US-08-352-678-6
32	35	59.3	230	4	PCT-US93-09636-6
33	35	59.3	259	2	US-07-857-224B-51
34	35	59.3	314	1	US-08-444-231-19
35	35	59.3	314	1	US-08-152-443A-19
36	35	59.3	314	4	PCT-US95-17083-4
37	35	59.3	335	2	US-08-219-237B-2
38	35	59.3	335	2	US-08-409-338-1
39	35	59.3	335	3	US-08-815-469-6
40	35	59.3	335	4	PCT-US95-17083-2
41	35	59.3	450	2	US-08-665-037-2
42	35	59.3	450	2	US-08-666-067-2
43	35	59.3	450	2	US-08-732-870-2
44	35	59.3	529	4	PCT-US95-05008-15
45	35	59.3	685	3	US-08-872-855-2

ALIGNMENTS

RESULT 1
US-08-794-796-2
; Sequence 2, Application US/08794796
; Patent No. 5885800
; GENERAL INFORMATION:
; APPLICANT: Emery, John
; APPLICANT: Tan, KB
; APPLICANT: Truneh, Alem
; APPLICANT: Young, Peter
; TITLE OF INVENTION: Tumor Necrosis Related Receptor,
; TITLE OF INVENTION: TR4
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/794,796
; FILING DATE: 04-FEB-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: GH50000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-4026
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 300 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-794-796-2

Query Match 100.0%; Score 59; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 0.06;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HASCPPGAGV 10
| | | | | | | | | |
Db 129 HASCPPGAGV 138

RESULT 2

US-08-974-022-4

; Sequence 4, Application US/08974022

; Patent No. 6015938

; GENERAL INFORMATION:

; APPLICANT: Boyle, William J.

; APPLICANT: Lacey, David L.

; APPLICANT: Calzone, Frank J.

; APPLICANT: Chang, Ming-Shi

; TITLE OF INVENTION: OSTEOPROTEGERIN

; NUMBER OF SEQUENCES: 53

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Amgen Inc.

; STREET: 1840 Dehavilland Drive

; CITY: Thousand Oaks

; STATE: California

; COUNTRY: USA

; ZIP: 91320-1789

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/974,022

; FILING DATE: 12-DEC-1995

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/577,788

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Winter, Robert B.

; REFERENCE/DOCKET NUMBER: A-378

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 401 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-974-022-4

Query Match 86.4%; Score 51; DB 3; Length 401;

Best Local Similarity 80.0%; Pred. No. 1.2;

Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 HASCPPGAGV 10
| | | | | | | | | |
Db 121 HRCPPGSGV 130

RESULT 3

US-08-974-022-2

; Sequence 2, Application US/08974022

; Patent No. 6015938

; GENERAL INFORMATION:

; APPLICANT: Boyle, William J.

; APPLICANT: Lacey, David L.

; APPLICANT: Calzone, Frank J.

; APPLICANT: Chang, Ming-Shi

; TITLE OF INVENTION: OSTEOPROTEGERIN

; NUMBER OF SEQUENCES: 53

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Amgen Inc.

; STREET: 1840 Dehavilland Drive

; CITY: Thousand Oaks

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-974-022-6

Query Match 81.4%; Score 48; DB 3; Length 401;
Best Local Similarity 80.0%; Pred. No. 3.4;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 HASCPPGAGV 10
Db 121 HRSCPPPGGV 130

RESULT 5
US-08-602-359A-34
; Sequence 34, Application US/08602359A
; Patent No. 5942430
; GENERAL INFORMATION:
; APPLICANT: ROBERTSON, Daniel E.
; APPLICANT: MURPHY, Dennis
; APPLICANT: REID, John
; APPLICANT: MAFFIA, Anthony
; APPLICANT: LINK, Steven
; APPLICANT: SWANSON, Ronald V.
; APPLICANT: WARREN, Patrick V.
; APPLICANT: KOSMOTKA, Anna
; TITLE OF INVENTION: ESTERASES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & RICHARDSON P.C.
; STREET: 4225 EXECUTIVE SQUARE, STE 1400
; CITY: LA JOLLA
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 92037

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/602,359A
FILING DATE: February 16, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: HAILE, LISA A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 09010/010001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-678-5070
TELEFAX: 619-678-5099
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 346 AMINO ACIDS
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-602-359A-34

Query Match 78.0%; Score 46; DB 2; Length 346;
Best Local Similarity 77.8%; Pred. No. 5.8;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 HASCPPGAG 9
Db 337 HECCPPGAG 345

RESULT 6
US-08-911-423-6
; Sequence 6, Application US/08911423
; Patent No. 611090
; GENERAL INFORMATION:
; APPLICANT: Gorman, Daniel M.
; APPLICANT: Randall, Troy D.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: MAMMALIAN CELL SURFACE ANTIGENS; RELATED
; TITLE OF INVENTION: REAGENTS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/911,423
FILING DATE: 14-AUG-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/023,419
FILING DATE: 16-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/027,901
FILING DATE: 07-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: DX0612K
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-852-9196
TELEFAX: 650-496-1200
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 228 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-911-423-6

Query Match 72.9%; Score 43; DB 3; Length 228;
Best Local Similarity 70.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 HASCPPGAGV 10
Db 85 HHCPPGGGV 94

RESULT 7
US-08-911-423-7
; Sequence 7, Application US/08911423
; Patent No. 611090
; GENERAL INFORMATION:
; APPLICANT: Gorman, Daniel M.
; APPLICANT: Randall, Troy D.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: MAMMALIAN CELL SURFACE ANTIGENS; RELATED
; TITLE OF INVENTION: REAGENTS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue

CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/911,423
FILING DATE: 14-AUG-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/023,419
FILING DATE: 16-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/027,901
FILING DATE: 07-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: DX0612K
TELEPHONE: 650-852-9196
TELEFAX: 650-496-1200
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 232 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-911-423-7

Query Match 72.9%; Score 43; DB 3; Length 232;
Best Local Similarity 70.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 HASCPPGAGV 10
| | | | |
Db 85 HHPCPPGGV 94

RESULT 8
US-08-911-423-4
; Sequence 4, Application US/08911423
; Patent No. 6111090
; GENERAL INFORMATION:
; APPLICANT: Gorman, Daniel M.
; APPLICANT: Randall, Troy D.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: MAMMALIAN CELL SURFACE ANTIGENS; RELATED
; TITLE OF INVENTION: REAGENTS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/911,423
; FILING DATE: 14-AUG-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/023,419
FILING DATE: 16-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/027,901
FILING DATE: 07-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: DX0612K
TELEPHONE: 650-852-9196
TELEFAX: 650-496-1200
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 241 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-911-423-4

Query Match 72.9%; Score 43; DB 3; Length 241;
Best Local Similarity 70.0%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 HASCPPGAGV 10
| | | | |
Db 91 HHPCPPGGV 100

RESULT 9
US-08-911-423-8
; Sequence 8, Application US/08911423
; Patent No. 6111090
; GENERAL INFORMATION:
; APPLICANT: Gorman, Daniel M.
; APPLICANT: Randall, Troy D.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: MAMMALIAN CELL SURFACE ANTIGENS; RELATED
; TITLE OF INVENTION: REAGENTS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/911,423
; FILING DATE: 14-AUG-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/023,419
; FILING DATE: 16-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/027,901
; FILING DATE: 07-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: DX0612K
; TELEPHONE: 650-852-9196
; TELEFAX: 650-496-1200
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 311 amino acids

; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 US-08-911-423-8

Query Match 72.9%; Score 43; DB 3; Length 311;
 Best Local Similarity 70.0%; Pred. No. 15;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 HASCPPGAGV 10
 Db 85 HPCPPGGV 94

RESULT 10

US-08-050-319B-39
 ; Sequence 39, Application US/08050319B
 ; Patent No. 5633145

GENERAL INFORMATION:

; APPLICANT: M.Feldmann, P.W. Gray,
 ; APPLICANT: M.J.C. Turner, F.M Brennan
 ; TITLE OF INVENTION: Modified human TNFalpha (Tumor
 ; TITLE OF INVENTION: Necrosis Factor alpha) Receptor
 ; NUMBER OF SEQUENCES: 57

CORRESPONDENCE ADDRESS:

; ADDRESSEE: Reed & Robbins
 ; STREET: 635 Bryant Street
 ; CITY: Palo Alto
 ; STATE: California
 ; COUNTRY: USA

ZIP: 94301

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/050,319B

FILING DATE: 10-May-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Robbins, Roberta L.

REGISTRATION NUMBER: 33,208

REFERENCE/DOCKET NUMBER: 5150-0030

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 617-8999

TELEFAX: (415) 327-3231

INFORMATION FOR SEQ ID NO: 39:

SEQUENCE CHARACTERISTICS:

LENGTH: 41 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-050-319B-39

Query Match 67.8%; Score 40; DB 1; Length 41;
 Best Local Similarity 70.0%; Pred. No. 6.1;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 HASCPPGAGV 10
 Db 19 HRSCSPGFGV 28

RESULT 11

US-08-465-982-39

Sequence 39, Application US/08465982

Patent No. 5863786

GENERAL INFORMATION:

; APPLICANT: M.Feldmann, P.W. Gray,
 ; APPLICANT: M.J.C. Turner, F.M Brennan
 ; TITLE OF INVENTION: Modified human TNFalpha (Tumor
 ; TITLE OF INVENTION: Necrosis Factor alpha) Receptor
 ; NUMBER OF SEQUENCES: 57

CORRESPONDENCE ADDRESS:

; ADDRESSEE: Reed & Robbins
 ; STREET: 635 Bryant Street
 ; CITY: Palo Alto
 ; STATE: California
 ; COUNTRY: USA

ZIP: 94301

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/465,982

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/050,319

FILING DATE: 10-May-1993

ATTORNEY/AGENT INFORMATION:

NAME: Robbins, Roberta L.

REGISTRATION NUMBER: 33,208

REFERENCE/DOCKET NUMBER: 5150-0030

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 617-8999

TELEFAX: (415) 327-3231

INFORMATION FOR SEQ ID NO: 39:

SEQUENCE CHARACTERISTICS:

LENGTH: 41 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-465-982-39

Query Match 67.8%; Score 40; DB 2; Length 41;
 Best Local Similarity 70.0%; Pred. No. 6.1;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 HASCPPGAGV 10
 Db 19 HRSCSPGFGV 28

RESULT 12

US-08-219-237B-7

Sequence 7, Application US/08219237B

Patent No. 5874546

GENERAL INFORMATION:

APPLICANT: NAGATA, Shigekazu

APPLICANT: ITOH, Naoto

APPLICANT: YONEHARA, Shin

TITLE OF INVENTION: DNA Coding for Human Cell Surface Antigen

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: James W. Hellwege

STREET: P.O. Box 2266 Eads Station

CITY: Arlington

STATE: Virginia

COUNTRY: USA

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/219,237B
FILING DATE: 28-MAR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/872,129
FILING DATE: 22-APR-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: James W. Hellwege
REGISTRATION NUMBER: 28,808
REFERENCE/DOCKET NUMBER: 516762
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 162 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-219-237B-7

Query Match 67.8%; Score 40; DB 2; Length 162;
Best Local Similarity 70.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 HASCPPGAV 10
| | | | | | | |
DB 98 HRSCSPFGV 107

RESULT 13
US-08-347-594A-2
; Sequence 2, Application US/08347594A
; Patent No. 5849536
; GENERAL INFORMATION:
; APPLICANT: Garfinkel, Leonard
; APPLICANT: Richter, Tamar
; TITLE OF INVENTION: CLONING AND PRODUCTION OF HUMAN VON
; TITLE OF INVENTION: WILLEBRAND FACTOR GPIB BINDING DOMAIN POLYPEPTIDES AND
; TITLE OF INVENTION: METHODS OF USING SAME
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/347,594A
; FILING DATE: No. 5849536ember 30, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36537-B2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2050 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-347-594A-2

Query Match 67.8%; Score 40; DB 2; Length 2050;
Best Local Similarity 85.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 HASCPPG 7
| | | | | | | |
DB 411 HAHCPPG 417
RESULT 14
US-08-896-449A-2
; Sequence 2, Application US/08896449A
; Patent No. 6040143
; GENERAL INFORMATION:
; APPLICANT: Venta, Patrick J
; APPLICANT: Yuzbasiyan-Gurkan, Vilma
; APPLICANT: Schall, William D
; APPLICANT: Brewer, George J
; TITLE OF INVENTION: DNA ENCODING CANINE VON WILLEBRAND
; TITLE OF INVENTION: FACTOR AND METHODS OF USE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
; STREET: 5445 Corporate Drive
; CITY: Troy
; STATE: Michigan
; COUNTRY: USA
; ZIP: 48098
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/896,449A
; FILING DATE: 18-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, DeAnn F.
; REFERENCE/DOCKET NUMBER: 2115-001226
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 248-641-1600
; TELEFAX: 248-641-0270
; TELEX: 287637
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2813 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-896-449A-2

Query Match 67.8%; Score 40; DB 3; Length 2813;
Best Local Similarity 85.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HASCPPG 7
| | | | | | | |
DB 1174 HAHCPPG 1180

RESULT 15
US-09-132-652-2
; Sequence 2, Application US/09132652
; Patent No. 6074832
; GENERAL INFORMATION:
; APPLICANT: Venta, Patrick J
; APPLICANT: Yuzbasiyan-Gurkan, Vilma
; APPLICANT: Schall, William D
; APPLICANT: Brewer, George J
; APPLICANT: Duffendeck, John
; TITLE OF INVENTION: DNA ENCODING CANINE VON WILLEBRAND FACTOR AND METHODS

; TITLE OF INVENTION: OF USE
; FILE REFERENCE: 2115S-001226CPB
; CURRENT APPLICATION NUMBER: US/09/132,652
; CURRENT FILING DATE: 1998-08-11
; EARLIER APPLICATION NUMBER: 08/896,449
; EARLIER FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2813
; TYPE: PRT
; ORGANISM: Canis familiaris
US-09-132-652-2

Query Match 67.8%; Score 40; DB 3; Length 2813;
Best Local Similarity 85.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 HASCPPG 7
|||
Db 1174 HAHCPPG 1180

Search completed: January 30, 2001, 16:47:52
Job time: 628 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:45:31 ; Search time 183.67 Seconds
(without alignments)
4.654 Million cell updates/sec

Title: US-09-518-931-4_COPY_142_166

Perfect score: 141

Sequence: 1 GESWARGAPRRGRRRCRGVAGP 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_36:*

- 1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT:*
- 2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT:*
- 3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT:*
- 4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT:*
- 5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT:*
- 6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT:*
- 7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT:*
- 8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT:*
- 9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT:*
- 10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT:*
- 11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT:*
- 12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT:*
- 13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT:*
- 14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT:*
- 15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT:*
- 16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT:*
- 17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT:*
- 18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT:*
- 19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT:*
- 20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT:*
- 21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	141	100.0	170	19	Human tumour necro
2	55.5	39.4	201	19	Human neurofilamen
3	55.5	39.4	386	18	Fragmented human N
4	55	39.0	416	19	Homo sapiens B115
5	51	36.2	125	20	Mycobacterium spec
6	51	36.2	128	20	Mycobacterium spec
7	51	36.2	470	20	Human Y1E ligand N
8	50.5	35.8	81	20	Human ovarian tumo
9	50.5	35.8	594	19	Mouse neuronal PAS
10	50	35.5	494	21	Amino acid sequenc
11	50	35.5	494	21	Amino acid sequenc
12	49	34.8	59	19	Human HUPF-I mutan

13	49	34.8	203	20	W97355	A partial HQB059
14	49	34.8	217	20	W97318	A HQB059 polypept
15	49	34.8	287	21	Y92704	Human IP-2 (N1K1 1
16	48.5	34.4	523	18	W22776	Human septin-2 pro
17	48.5	34.4	674	20	Y29125	Amino acid sequenc
18	48	34.0	40	19	Y21529	WO9845322 Seq ID 1
19	48	34.0	300	18	W18661	Fragmented human N
20	48	34.0	637	20	Y04993	Mycobacterium spec
21	48	34.0	1184	14	R32882	Cardiac adenylyl c
22	48	34.0	1184	16	R78519	Cardiac adenylyl c
23	48	34.0	5069	19	W52846	A. mediterranei r1
24	47.5	33.7	218	18	W30839	MyoK1 protein myok
25	47.5	33.7	412	14	R41525	Human A2a adenosin
26	47.5	33.7	412	15	R47387	Human A2a adenosin
27	47.5	33.7	412	16	R93990	Human ventricule A2
28	47.5	33.7	412	16	R87656	Human adenosine re
29	47.5	33.7	412	16	R84193	Human A2a adenosin
30	47.5	33.7	412	20	W82312	Human A2a adenosin
31	47	33.3	228	20	Y00347	Fragment of human
32	47	33.3	275	21	Y74422	Neisseria meningit
33	47	33.3	276	21	Y74421	Neisseria meningit
34	47	33.3	386	21	Y82725	Thiosulphate sulph
35	47	33.3	983	20	Y09513	Mouse JMY protein
36	46.5	33.0	69	20	Y19742	SEQ ID NO 460 from
37	46.5	33.0	298	20	W98080	Rye chitinase-like
38	46.5	33.0	318	20	W98079	Rye chitinase-like
39	46.5	33.0	1958	15	R60620	Protein from ORF2
40	46	32.6	101	20	W30654	Human secreted pro
41	46	32.6	115	20	Y74062	Human prostate tum
42	46	32.6	117	18	W36147	Rat eosinophil-der
43	46	32.6	169	19	Y21034	Human glial fibril
44	46	32.6	183	21	Y75341	Neisseria meningit
45	46	32.6	227	18	W36148	Rat eosinophil-der

ALIGNMENTS

RESULT 1

W63623
ID W63623 standard; Protein; 170 AA.

XX W63623;

XX 26-OCT-1998 (first entry)

XX Human tumour necrosis factor receptor-6 beta protein.

XX Human tumour necrosis factor receptor-6 alpha; TNFR-6 alpha; TNFR-6 beta; endothelial cells; keratinocytes; normal prostate; apoptosis; prostate tumour tissue.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..30

FT Protein 31..170

FT note= "TNFR-6 beta"

FT Region 31..166

FT /note= "Soluble extracellular domain"

XX WO9830694-A2.

XX 16-JUL-1998.

XX 13-JAN-1998; 98WO-US00153.

XX 14-JAN-1997; 97US-0035496.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ebner R, Feng P, Gentz RL, Ni J, Ruben SM, Yu G;

DR WPI; 1998-399142/34.
 DR N-PSDB; V39086.
 XX
 XX Human tumour necrosis factor receptors 6-alpha and 6-beta - used in
 PT the diagnosis of immune system-related disorder(s)
 XX
 PS Claim 20; Fig 2A-2B; 91pp; English.
 XX
 XX The present sequence represents the human tumour necrosis factor
 CC receptor-6 beta (TNFR-6 beta) protein. The invention also provides
 CC for the TNFR-6 alpha protein (W63622). TNFR-6 alpha and TNFR-6 beta are
 CC members of the tumour necrosis factor receptor (TNFR) family. TNFRs
 CC are expressed in endothelial cells, keratinocytes, normal prostate and
 CC prostate tumour tissue. For a number of disorders of these cells,
 CC particularly of the immune system, substantially altered (whether
 CC increased or decreased) levels of TNFR-6 alpha and/or TNFR-6 beta
 CC expression can be detected, therefore the TNFR-6 alpha and TNFR-6 beta
 CC polypeptides, nucleic acids and antibodies are claimed to be useful in
 CC the diagnosis of such disorders. Mutations of the TNFR-6 alpha and
 CC TNFR-6 beta genes can also be detected. The TNFR polypeptides are
 CC also claimed to be useful for identifying ligands which may be useful
 CC in the treatment of apoptosis related disorders.
 XX
 SQ Sequence 170 AA;
 Query Match 100.0%; Score 141; DB 19; Length 170;
 Best Local Similarity 100.0%; Pred. No. 1.6e-11;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GESWARGAPRSGRRRCGRGVAGP 25
 DB 142 geswargaprsgrgrgrgvagp 166
 |||||
 RESULT 2
 Y20852
 ID Y20852 standard; Protein: 201 AA.
 XX
 AC Y20852;
 XX
 DT 22-JUL-1999 (first entry)
 XX
 DE Human neurofilament-H mutant protein fragment 11.
 XX
 KW Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
 KW frameshift mutation; age-related disease; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPP-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO9845322-A2.
 XX
 PD 15-OCT-1998.
 XX
 PF 02-APR-1998; 98WO-IB00705.
 XX
 PR 10-APR-1997; 97US-0043163.
 XX
 PA (UYUT-) RIJKSUNIV UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.
 XX
 PI Burbach JPH, Grosveld FG, Van Leeuwen FW;
 XX

DR WPI; 1998-609901/51.
 DR N-PSDB; X75760.
 XX
 XX Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA
 XX
 PS Disclosure; Figure 9; 258pp; English.
 XX
 CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPP-1, high mobility group
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.
 XX
 SQ Sequence 201 AA;
 Query Match 39.4%; Score 55.5; DB 19; Length 201;
 Best Local Similarity 57.1%; Pred. No. 3;
 Matches 12; Conservative 1; Mismatches 7; Indels 1; Gaps 1;
 QY 6 RRGAPRSGRRRC-GRGVAGP 25
 DB 119 rggpraggrgrgrgargp 139
 |||||
 RESULT 3
 W18664
 ID W18664 standard; Protein: 386 AA.
 XX
 AC W18664;
 XX
 DT 24-JUL-1997 (first entry)
 XX
 DE Fragmented human NF-H gene +1 frameshift mutant product.
 XX
 KW Frameshift mutation product; GAGA motif; somatic mutation; diagnosis;
 KW detection; antibody; probe; cancer; neoplasia; neurodegenerative;
 KW Parkinson's; Alzheimer's disease; Pick's; Huntington's disease;
 KW Down's syndrome; frontal lobe dementia; progressive supranuclear palsy;
 KW PSP; amyotrophic lateral sclerosis; multiple sclerosis; MS;
 KW cardiovascular; rheumatoid arthritis.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1..387
 FT /note= "X corresponds to a stop codon in the
 FT accompanying DNA file, T69796"
 FT
 FT Peptide 129..135
 FT /note= "antigenic peptide used fro antibody
 FT production"
 FT
 XX WO9712992-A2.
 XX
 PD 10-APR-1997.
 XX

PF 02-OCT-1996; 96WO-IB011106.
 XX 11-JAN-1996; 96US-0009832.
 PR 02-OCT-1995; 95GB-0020080.
 XX (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.
 PA (UYUT-) UNIV STATE UTRECHT.
 XX Burchach JPH, Grosveld FG, Van Leeuwen FW;
 PI N-PSDB; T69796.
 XX WPI; 1997-226235/20.
 DR N-PSDB; T69796.
 XX Use of mutant genes having frame:shift mutation(s) - for developing
 PT prods. for the diagnosis, prevention and treatment of associated
 PT diseases, e.g. cancer or neurodegenerative disease
 XX Cialm 22; Fig 9; 123pp; English.
 XX W18663 and W18664 are +2 and +1 frameshift mutations, respectively,
 CC of a sequence comprising fragments of the coding sequence of the
 CC human neurofilament subunit NF-H gene corresponding to nucleotides
 CC 1-1162 of the wild-type NF-H gene. This region contains GAGAG motifs.
 CC Frameshift mutants of the tau, ubiquitin, apolipoprotein E,
 CC microtubule-associated protein 2 (MAP-2), neurofilament subunit L, M
 CC and H and amyloid A4 genes are claimed. All these genes share a common
 CC GAGAN motif (N- A, G, C or T), which is the site of common GA
 CC dinucleotide deletions(s) that cause neurodegenerative disorders.
 CC Antigenic peptides used for the production of antibodies, and small
 CC nucleic acid sequences derived from frameshift mutants are used in the
 CC diagnosis, prevention and treatment of cancer and neurodegenerative
 CC diseases, e.g. Parkinson's disease, Alzheimer's disease, Down's
 CC syndrome, frontal lobe dementia (Pick's disease), progressive
 CC supranuclear palsy (PSP), amyotrophic lateral sclerosis, Huntington's
 CC disease, multiple sclerosis, and other degenerative diseases such as
 CC cardiovascular disease and rheumatoid arthritis.
 XX Sequence 386 AA;
 SQ

Query Match 39.4%; Score 55.5; DB 18; Length 386;
 Best Local Similarity 57.1%; Pred. No. 5.7;
 Matches 12; Conservative 1; Mismatches 7; Indels 1; Gaps 1;
 QY 6 RGGAPRSGGRC-CRGQVAGP 25
 ||| ||| | ||| ||
 Db 288 rgggpragralrgrgargp 308

RESULT 4
 W54098
 ID W54098 standard; Protein; 416 AA.
 XX AC
 XX W54098;
 XX 28-SEP-1998 (first entry)
 DE Homo sapiens B115 sequence.
 XX BARD1; ring protein; BRCA1; breast cancer; risk; diagnosis.
 KW Homo sapiens.
 OS Homo sapiens.
 XX W09812327-A2.
 PN W09812327-A2.
 XX 26-MAR-1998.
 PD 19-SEP-1997; 97WO-US16842.
 PF 04-APR-1997; 97US-0042985.
 XX 20-SEP-1996; 96US-0025296.
 PR 03-APR-1997; 97US-0042611.

XX (TEXA) UNIV TEXAS SYSTEM.
 PA Baer R, Bowcock AM;
 PI WPI; 1998-230317/20.
 XX N-PSDB; V24134.
 DR DNA sequence encoding BARD1, B123, BE2, BE14, BE31 or BE445 - which
 PT as breast cancer antigen, BRCA1, binding proteins are useful to
 PT identify patient having or at risk of developing cancer
 XX Disclosure; Page 285-286; 348pp; English.
 XX The sequence is that of a protein which can be used in the
 CC preparation of the recombinant breast cancer antigen, BRCA1, binding
 CC proteins BARD1, B123, BE14, BE31 or BE445, or a composition for the
 CC detection of a BARD1, B123, BE2, BE14, BE31 or BE445 nucleic acid
 CC sequence, specifically a wild type BARD1 composition for the detection
 CC or purification of BRCA1, useful to identify a patient having, or at
 CC risk of developing cancer. BARD1 can be used in the preparation of an
 CC anti-BARD1 antibody, and in the detection and purification of a BRCA1
 CC protein. BARD1, B123, BE2, BE14, BE31 or BE445 can be used in the
 CC identification of a binding protein agonist or antagonist that alters
 CC the binding of BARD1, B123, BE2, BE14, BE31 or BE445 to BRCA1 or the
 CC biological activity of the BRCA1-BARD1, B123, BE2, BE14, BE31 or BE445
 CC complex. The antibodies can be used to detect BARD1, B123, BE2, BE14,
 CC BE31 or BE445, a specific anti-BARD1 antibody can be used to identify
 CC a patient having or at risk of developing cancer.
 XX Sequence 416 AA;
 SQ

Query Match 39.0%; Score 55; DB 19; Length 416;
 Best Local Similarity 43.5%; Pred. No. 7.1;
 Matches 10; Conservative 2; Mismatches 11; Indels 0; Gaps 0;
 QY 2 ESWARGGAPRSGRCGRGQVAG 24
 :|| ||| ||| ||| |||
 Db 4 qnwgdrgrssrggygggpagg 26

RESULT 5
 Y04858
 ID Y04858 standard; Protein; 125 AA.
 XX AC Y04858;
 XX 06-JUL-1999 (first entry)
 DT Mycobacterium species protein sequence 18C'.
 DE Secreted protein; Mycobacterium; primer: PCR; amplification: probe;
 KW hybridisation; detection; vaccine; immunisation; infection.
 XX Mycobacterium sp.
 OS Mycobacterium sp.
 PN W09909186-A2.
 XX 25-FEB-1999.
 PD 14-AUG-1998; 98WO-FR01813.
 XX 11-SEP-1997; 97FR-0011325.
 PR 14-AUG-1997; 97FR-0010404.
 XX (INSP) INST PASTEUR.
 PA Gicquel B, Lim EM, Pelicic V, Portnoi D, Coguet de la Salmoniere Y;
 PI Guigueno A;
 XX WPI; 1999-181045/15.
 DR N-PSDB; X34110.

XX Mycobacterial DNA vectors containing reporter constructs - for
PT identifying coding or promoter sequences involved in
PT infection-associated protein expression

XX
PS Claim 32; Fig 18C'; 309pp; French.

XX
CC Sequences Y04742-Y05000 and Y07201-Y07204 represent secreted proteins
CC from various Mycobacterium species microorganisms. The encoding
CC nucleotide sequences can be used as primers and probes for methods
CC for detecting and identifying mycobacteria, especially belonging to
CC the M. tuberculosis complex. The encoded proteins can be used in
CC vaccines for immunisation against a bacterial or viral infection.

XX
SQ Sequence 125 AA;

Query Match 36.2%; Score 51; DB 20; Length 125;
Best Local Similarity 64.3%; Pred. No. 7.4;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 6 RGGAPRSGGRRRCGR 19
|| | : | |||||
Db 66 rgaepatgprrcgr 79

RESULT 6
Y04854 Y04854 standard; Protein; 128 AA.
XX AC Y04854;
XX
XX 06-JUL-1999 (first entry)
XX
DE Mycobacterium species protein sequence 18B.
XX
KW Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
KW hybridisation; detection; vaccine; immunisation; infection.
XX
OS Mycobacterium sp.
XX
XX W09909186-A2.
XX
XX 25-FEB-1999.
XX
XX 14-AUG-1998; 98WO-FR01813.
XX
XX 11-SEP-1997; 97FR-0011325.
XX
XX 14-AUG-1997; 97FR-0010404.
XX
XX (INSP) INST PASTEUR.
XX
XX Gicquel B, Lim EM, Pellicic V, Portnoi D, Goguet de la Salmoniere Y;
XX Guigueno A;
XX
XX WPI; 1999-181045/15.
XX
XX N-PSDB; X34106.
XX
XX Mycobacterial DNA vectors containing reporter constructs - for
PT identifying coding or promoter sequences involved in
PT infection-associated protein expression

XX
PS Claim 32; Fig 18B; 309pp; French.

XX
CC Sequences Y04742-Y05000 and Y07201-Y07204 represent secreted proteins
CC from various Mycobacterium species microorganisms. The encoding
CC nucleotide sequences can be used as primers and probes for methods
CC for detecting and identifying mycobacteria, especially belonging to
CC the M. tuberculosis complex. The encoded proteins can be used in
CC vaccines for immunisation against a bacterial or viral infection.

XX
SQ Sequence 128 AA;

Query Match 36.2%; Score 51; DB 20; Length 128;
Best Local Similarity 64.3%; Pred. No. 7.6;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 6 RGGAPRSGGRRRCGR 19
|| | : | |||||
Db 70 rgaepatgprrcgr 83

RESULT 7
Y05398 Y05398 standard; Protein; 470 AA.
XX AC Y05398;
XX
XX 01-JUL-1999 (first entry)
XX
DE Human TIE ligand NL8 protein sequence.
XX
XX Human TIE ligand; NL1; NL4; NL5; NL8; Ig homology domain; angiogenesis;
KW EGF homology domain; receptor tyrosine kinase; vascular endothelial cell;
KW early haemopoietic cell; haemopoiesis; neovascularisation; wound healing;
KW endothelial cell; growth inhibitor; apoptosis inducer; tumour cell;
KW vasculogenesis; detection; diagnosis; therapy.
XX
XX Homo sapiens.
XX
XX W09915653-A2.
XX
XX 01-APR-1999.
XX
XX 14-SEP-1998; 98WO-US19093.
XX
XX 29-OCT-1997; 97US-0960507.
XX
XX 19-SEP-1997; 97US-0933821.
XX
XX (GETH) GENENTECH INC.
XX
XX Botstein D, Ferrara N, Goddard A, Godowski PJ, Gurney AL;
XX Hillan K, Roy M, Schwall R, Tumas D;
XX WPI; 1999-263480/22.
XX
XX N-PSDB; X36343.
XX
XX New isolated TIE ligand homologs for, e.g. developing products for
PT treatment of tumors
XX
XX Claim 8; Fig 7; 132pp; English.

XX
CC This sequence is the human tyrosine kinase containing Ig and EGF
CC homology domains (TIE) ligand of the invention, designated NL8.
CC The TIE receptors are receptor tyrosine kinases which are expressed in
CC vascular endothelial cells and early haemopoietic cells. The TIE
CC receptors are believed to be actively involved in angiogenesis, and may
CC play a role in haemopoiesis as well. The TIE ligand homologs can promote
CC the survival and/or growth and/or differentiation of TIE receptor
CC expressing cells. They can be used for promoting neovascularisation in
CC wound healing and for promoting angiogenic processes, such as for
CC inducing collateral vascularisation in an ischaemic heart or limb, or for
CC promoting bone development and/or maturation and/or growth in a patient
CC or muscle growth and development. The TIE ligand homologs and antibodies
CC can inhibit the growth of endothelial cells and induce apoptosis of
CC cells, particularly tumour cells. They can inhibit vasculogenesis,
CC particularly the vascularisation of tumour cells. The antibodies can also
CC inhibit vascularisation of a cell in which a gene encoding an NL1, NL5,
CC NL8 or NL4 polypeptide is amplified. The products can also be used for
CC detection, diagnosis, drug screening and production of transgenic
CC animals.

XX
SQ Sequence 470 AA;

Query Match 36.2%; Score 51; DB 20; Length 470;
Best Local Similarity 81.8%; Pred. No. 27;
Matches 9; Conservative 0; Mismatches 2; Indels

Qy	1	GESWARGGAPR	11
Db	16	gaswaragapr	26

```

RESULT      8
Y76537
ID          Y76537 standard; Protein; 81 AA.
XX
XX
AC          Y76537;
XX
XX
DT          10-APR-2000 (first entry)
XX
DE          Human ovarian tumor EST fragment encoded protein 33.
XX
XX
KW          Expressed sequence tag; EST; human; ovarian tumor; anticancer;
KW          gene therapy; treatment.

```

Query Match 35.8%; Score 50.5; DB 20; Length 81;
Best Local Similarity 60.0%; Pred. No. 5.7;
Matches 12: Conservative 2; Mismatches 5; Indels 1; Gaps 1;

Qy 7 GGAP-RSGGRRCGRGVAGP 25
||| | | | | : : |
Db 29 qgaprsrqrriqwarvssp 48

RESULT	9	
W68092		
ID	W68092 standard; Protein: 594 AA.	
XX		
XX	W68092;	
XX		
DT	09-NOV-1998 (first entry)	
XX		
DE	Mouse neuronal PAS domain protein NPAS1.	
XX		
KW	NPAS1; neuronal bHLH-PAS domain; mouse; tran-	
KW	therapy; diagnosis.	
XX		
OS	Mus sp.	
XX		
XX		
FH	Key	Location/Qualifiers
FT	Domain	1..134
ET		/note= "NPAS protein domain"
XX		
XX	WO9831804-A1.	
PN		
XX		
PD	23-JUL-1998.	
XX		
XX		
PF	21-JAN-1998; 98WO-US01154.	
XX		
XX		
PR	21-JAN-1997; 97US-0785310.	
XX		
XX	(TEXA) UNIV TEXAS SYSTEM.	
PA		
XX		
PI	McKnight SL, Russel DW;	
XX		
DR	WPI; 1998-414103/35.	
DR	N-PSDB; V41257.	
XX		
XX		
PT	New isolated neuronal PAS domain proteins -	
PT	neurological tissue such as brain tissue, us	
PT	for diagnosis and therapy	
XX		
PS	Claim 1; Page 29-31; 42pp: English	

Query Match	35.8%	Score 50.5;	DB 19;	Length 594;
Best Local Similarity	61.1%	Pred. No. 39;		
Matches 11: Conservative	2;	Mismatches	2;	Indels 3;
				Gaps 1;

```
QY      10 PRSGR---RCGRQVAG 24
          ||||| :||| :||
Db      6 prsgrgvevkcggrag 23
```

RESULT 10

Y69386
ID Y69386 standard; Protein; 494 AA.
XX
AC Y69386;
XX
DT 19-JUN-2000 (first entry)
XX
DE Amino acid sequence of a human OC-3 protein.
XX
KW Human; OC-3; ONECUT protein; HNF-6; islets of Langerhans;
KW insulin response; diabetes; cancer; melanoma; Waardenburg syndrome;
KW cellular therapy; artificial pancreas.
XX
OS Homo sapiens.
XX
PN WO200011159-A1.
XX
PD 02-MAR-2000.
XX
PF 17-AUG-1999; 99WO-BE00112.
XX
PR 17-AUG-1998; 98BE-0000609.
XX
PA (UYLO-) UNIV CATHOLIQUE LOUVAIN.
XX
PI Rousseau G, Lemaigre F;
XX
DR WPI; 2000-224694/19.
DR N-PSDB; Z61590.
XX
XX New composition useful for treating or preventing diabetes and cancer,
PT contains nucleic acid encoding a ONECUT family protein or corresponding
PT vector or transformed cell
XX
PS Disclosure; Page 30-32; 38pp; French.
XX
XX The present sequence represents a human protein (OC-3) which is a member
CC of the ONECUT family. The ONECUT family protein HNF-6 is essential for
CC formation of the islets of Langerhans and for insulin responses.
CC ONECUT polypeptides and polynucleotides are used to treat and/or
CC prevent diabetes types 1 and 2 and associated disorders, cancer,
CC particularly melanoma and Waardenburg syndrome. The polypeptide is
CC especially useful in cellular therapy to create an artificial
CC pancreas.
XX
SQ Sequence 494 AA;

Query Match 35.5%; Score 50; DB 21; Length 494;
Best Local Similarity 44.0%; Pred. No. 38;
Matches 11; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 1 GESWARGGAPRSGRCRGQVAGP 25
Db 63 ggaggaggagsgaggadfrgelagg 87
| : ||| : || ||| : |||
| : ||| : || ||| : |||

RESULT 11
Y69393
ID Y69393 standard; Protein; 494 AA.
XX
AC Y69393;
XX
DT 19-JUN-2000 (first entry)
XX
DE Amino acid sequence of a human OC-3 protein.
XX
KW Human; OC-3; ONECUT protein; transgenic animal; diabetes model;
KW cancer.
XX
OS Homo sapiens.
XX
PN WO200011158-A1.

XX
PD 02-MAR-2000.
XX
PF 13-AUG-1999; 99WO-BE00108.
XX
PR 17-AUG-1998; 98BE-0000608.
XX
PA (UYLO-) UNIV CATHOLIQUE LOUVAIN.
XX
PI Rousseau G, Lemaigre F;
XX
DR WPI; 2000-224693/19.
DR N-PSDB; Z61593.
XX
XX New transgenic animal, useful as model for diabetes and cancer
PT development, has a deletion in a gene that encodes a protein of the
PT ONECUT family
XX
PS Disclosure; Page 24-26; 33pp; French.
XX
XX The present sequence represents a human protein (OC-3) that is a
CC member of the ONECUT family. The specification describes a non-human,
CC genetically modified animal which has a partial or complete deletion
CC in a ONECUT gene. The transgenic animals are useful as models for
CC diabetes and/or cancer development. The animals are especially for
CC testing therapeutic or prophylactic procedures or agents, and for
CC precise analysis of physiological and biological processes involved
CC in the development of these diseases.
XX
SQ Sequence 494 AA;

Query Match 35.5%; Score 50; DB 21; Length 494;
Best Local Similarity 44.0%; Pred. No. 38;
Matches 11; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 1 GESWARGGAPRSGRCRGQVAGP 25
Db 63 ggaggaggagsgaggadfrgelagg 87
| : ||| : || ||| : |||
| : ||| : || ||| : |||

RESULT 12
Y21350
ID Y21350 standard; Protein; 59 AA.
XX
AC Y21350;
XX
DT 22-JUL-1999 (first entry)
XX
DE Human HUPF-I mutant protein fragment 2.
XX
KW Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
KW frameshift mutation; age-related disease; neurodegenerative disorder;
KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
KW high mobility group protein-C; neuroendocrine specific protein A.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9845322-A2.
XX
PD 15-OCT-1998.
XX
PF 02-APR-1998; 98WO-IB00705.
XX
PR 10-APR-1997; 97US-0043163.
XX

PA (UYUT-) RIJKSUNIV UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 PA (UVRO-) UNIV ROTTERDAM ERASMUS.
 XX
 XX
 PI Burbach JPH, Grosveld FG, Van Leeuwen FW;
 XX
 DR WPI; 1998-609901/51.
 DR N-PSDB; X75768.
 XX
 XX
 PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA
 XX
 XX
 PS Disclosure; Figure 17; 258pp; English.
 XX
 CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-ApP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPP-1, high mobility group
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.
 XX
 SQ Sequence 59 AA;

Query Match 34.8%; Score 49; DB 19; Length 59;
 Best Local Similarity 60.0%; Pred. No. 6.6;
 Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 4 WARGAPRSGRRRCG 18
 | | | | |
 Db 15 wrkgpgrgrircg 29

RESULT 13
 W97355
 ID W97355 standard; Protein; 203 AA.
 XX
 AC W97355;

DT 11-MAY-1999 (first entry)
 XX
 DE A partial HOQBQ59 polypeptide sequence.

XX HOQBQ59 polypeptide; bone loss; osteoporosis; inflammatory disease;
 KW HIV-2; viral infection; immunodeficiency disorder; septic shock; pain;
 KW injury; cancer; anorexia; bulimia; Parkinson's disease;
 KW cardiovascular disease; urinary retention; angina pectoris; ulcer;
 KW benign prostatic hypertrophy; neurological disorder;
 KW severe mental retardation; dyskinesias; vaccine.

OS Homo sapiens.
 XX
 XX WO9902560-A1.
 XX
 XX 21-JAN-1999.
 PD
 XX 09-JUL-1998; 98WO-US14291.
 PF
 XX

PR 13-NOV-1997; 97US-0969239.
 PR 09-JUL-1997; 97US-0052009.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX
 PI DeMarini DJ;
 XX
 DR WPI; 1999-120784/10.
 DR N-PSDB; X15871.
 XX
 XX
 PT New HOQBQ59 polypeptide - useful for diagnosing or treating bone
 PT loss, inflammatory and other immunodeficiency diseases
 XX
 XX
 PS Disclosure; Page 14; 41pp; English.

XX The present sequence represents a partial HOQBQ59 polypeptide. The
 CC protein can be used in diagnostic assays to detect diseases associated
 CC with inappropriate HOQBQ59 activity or levels. It is also useful for
 CC chromosome assays to detect mutations in the gene. Diseases which may
 CC be diagnosed or treated are bone loss including osteoporosis;
 CC inflammatory diseases such as Adult Respiratory Disease Syndrome (ARDS),
 CC rheumatoid arthritis, osteoarthritis, inflammatory bowel disease (IBD),
 CC psoriasis, dermatitis, asthma, allergies; infections such as bacterial,
 CC fungal, protozoan and viral, particularly infections caused by HIV-1 or
 CC HIV-2; HIV-associated cachexia and other immunodeficiency disorders;
 CC septic shock; pain; injury; cancer including testicular cancer; anorexia;
 CC bulimia; Parkinson's disease; cardiovascular disease including
 CC restenosis, atherosclerosis, acute heart failure, myocardial infarction;
 CC hypotension; hypertension; urinary retention; angina pectoris; ulcers;
 CC benign prostatic hypertrophy; and psychotic and neurological disorders,
 CC including anxiety, schizophrenia, manic depression, delirium, dementia,
 CC severe mental retardation and dyskinesias, such as Huntington's disease
 CC or Gilles de la Tourette's syndrome. The protein may be used in
 CC vaccines.
 XX

SQ Sequence 203 AA;

Query Match 34.8%; Score 49; DB 20; Length 203;
 Best Local Similarity 46.4%; Pred. No. 22;
 Matches 13; Conservative 0; Mismatches 11; Indels 4; Gaps 1;

Oy 1 GESWARGAPRSGGR- ---RCGRGOVAG 24
 | | | | | | | | | |
 Db 88 grgggrgggrgggrgggrgggrggggg 115

RESULT 14
 W97318
 ID W97318 standard; Protein; 217 AA.

XX W97318;
 AC
 XX
 DT 11-MAY-1999 (first entry)
 XX
 DE A HOQBQ59 polypeptide sequence.

XX HOQBQ59 polypeptide; bone loss; osteoporosis; inflammatory disease;
 KW HIV-2; viral infection; immunodeficiency disorder; septic shock; pain;
 KW injury; cancer; anorexia; bulimia; Parkinson's disease;
 KW cardiovascular disease; urinary retention; angina pectoris; ulcer;
 KW benign prostatic hypertrophy; neurological disorder;
 KW severe mental retardation; dyskinesias; vaccine.

OS Homo sapiens.
 XX
 XX WO9902560-A1.
 XX
 XX 21-JAN-1999.
 PD
 XX 09-JUL-1998; 98WO-US14291.
 PF
 XX
 XX 13-NOV-1997; 97US-0969239.

```
PR 09-JUL-1997; 97US-0052009.
XX (SMIK ) SMITHLINE BEECHAM CORP.
XX DeMarini DJ;
XX WPI; 1999-120784/10.
DR N-PSDB; X15856.
XX
XX New HOQB059 polypeptide - useful for diagnosing or treating bone
PT loss, inflammatory and other immunodeficiency diseases
XX
XX Claim 2; Page 11-12; 41pp; English.
XX
XX The present sequence represents a HOQB059 polypeptide. The protein can
CC be used in diagnostic assays to detect diseases associated with
CC inappropriate HOQB059 activity or levels. It is also useful for
CC chromosome assays to detect mutations in the gene. Diseases which may
CC be diagnosed or treated are bone loss including osteoporosis;
CC inflammatory diseases such as Adult Respiratory Disease Syndrome (ARDS),
CC rheumatoid arthritis, osteoarthritis, inflammatory bowel disease (IBD),
CC psoriasis, dermatitis, asthma, allergies; infections such as bacterial,
CC fungal, protozoan and viral, particularly infections caused by HIV-1 or
CC HIV-2; HIV-associated cachexia and other immunodeficiency disorders;
CC septic shock; pain; injury; cancer including testicular cancer; anorexia;
CC bulimia; Parkinson's disease; cardiovascular disease including
CC restenosis, atherosclerosis, acute heart failure, myocardial infarction;
CC hypotension; hypertension; urinary retention; angina pectoris; ulcers;
CC benign prostatic hypertrophy; and psychotic and neurological disorders,
CC including anxiety, schizophrenia, manic depression, delirium, dementia,
CC or Gilles de la Tourette's syndrome. The protein may be used in
XX vaccines.
XX
XX Sequence 217 AA;
XX
XX Query Match 34.8%; Score 49; DB 20; Length 217;
XX Best Local Similarity 46.4%; Pred. No. 23;
XX Matches 13; Conservative 0; Mismatches 11; Indels 4; Gaps 1;
XX
XX QY 1 GESWARGGAPRSGR-----RCGRGQVAG 24
XX | ||| | ||| | ||| |
XX Db 175 ggggrgggrgggrgggrgggrgggrggg 202
XX
XX RESULT 15
XX Y92704
XX ID Y92704 standard; Protein: 287 AA.
XX AC
XX Y92704;
XX
XX 10-AUG-2000 (first entry)
XX
XX Human IP-2 (NIK1 interacting protein 2) (version 2).
XX
XX IP-2; NIK1 Interacting Protein; EST; aal43467; tropomyosin;
XX protein complex; cytostatic; antiviral; neuroprotective; cardiant.
XX
XX Homo sapiens.
XX
XX WO200020448-A2.
XX
XX 13-APR-2000.
XX
XX 06-OCT-1999; 99WO-US23314.
XX
XX 06-OCT-1998; 98US-0167206.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Nandabalan K, Schulz VP, Yang M;
XX
XX
```

```
DR WPI; 2000-303742/26.
DR N-PSDB; A09468.
XX
XX New complex of a NLK1 protein and a NLK1 protein-interacting protein,
PT useful for treating cancer, hyperproliferative disorder.
PT neurodegenerative disorder, cardiomyopathies, viral infections and
XX metabolic disorders
XX
XX Claim 43; Fig 10; 172pp; English.
XX
XX The DNA encoding IP-2 protein (NIK1 Interacting Protein) was identified
CC using a modified yeast two hybrid system with NIK1 as "bait". IP-2 DNA
CC was identical to EST aal43467. The IP-2 protein revealed homology to
CC alpha-tropomyosin chain from the skeletal muscle.
CC The invention concerns purified complexes of a NLK1 protein
CC and a NLK1 protein-interacting protein, where the interacting protein is
CC selected from TrkA, protein phosphatase laipha, 14-3-3epsilon,
CC alpha-tropomyosin, vimentin, p0071, Inl-1, IP-1, IP-2, IP-3, IP-4, or
CC IP-5. NIK1 (also referred to as Nek2) is a human homologue of the
CC Aspergillus nidulans mitotic regulator, NIMA kinase. NIK1 is a
CC serine/threonine-specific kinase and is thought to play a key role in
CC cell-cycle events leading to the onset of mitosis. The complexes, their
CC derivatives and NLK1 or NLK1-IP protein and DNA sequences, etc. are
CC useful for treating or preventing a disease or disorder involving
CC aberrant levels of the complex or protein. Such disorders include
CC cancer, hyperproliferative disorders, neurodegenerative disorders,
CC cardiomyopathies, viral infections and metabolic disorders.
CC Note: This sequence is referred to as SEQ. ID. 20 in the body of the
CC specification. A different sequence appears in the SEQ. ID. listing,
CC also called SEQ. ID. 20 (see Y92326).
XX
XX Sequence 287 AA;
XX
XX Query Match 34.8%; Score 49; DB 21; Length 287;
XX Best Local Similarity 41.4%; Pred. No. 30;
XX Matches 12; Conservative 1; Mismatches 10; Indels 6; Gaps 2;
XX
XX QY 3 SWARGG-APRSGG-----RRCGRGQVAG 25
XX :|| ||| ||| ||| |||
XX Db 3 awaeacgprgggagapracsrwgqegp 31
XX
XX Search completed: January 30, 2001, 16:45:32
XX Job time: 592 sec
```


THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:50:39 ; Search time 149.64 seconds
(without alignments)
11.344 Million cell updates/sec

Title: US-09-518-931-4_COPY_142_166

Perfect score: 141

Sequence: 1 GESWARGGAPRSGRGRCGRGVAGP 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_66: *
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	43.3	55	2 T36132	hypothetical prote
2	59	41.8	320	2 T09555	fibrillarlin - Arab
3	57.5	40.8	199	2 T36594	probable single-st
4	57.5	40.8	573	2 JQ0135	hypothetical 62.8K
5	56	39.7	210	2 B75450	hypothetical prote
6	54	38.3	152	2 C72662	hypothetical prote
7	54	38.3	168	2 S78110	thi protein - Rhiz
8	54	38.3	198	2 A57717	transcription fact
9	53	37.6	129	2 S14984	glycine-rich prote
10	53	37.6	889	2 T33422	hypothetical prote
11	52.5	37.2	166	2 B75308	ribonuclease P pro
12	52.5	37.2	886	2 S07132	hypothetical prote
13	52	36.9	305	2 S33690	fibrillarlin - fliss
14	52	36.9	327	2 S25421	nucleolar protein
15	52	36.9	340	2 C70743	hypothetical prote
16	51	36.2	71	1 I40863	piscicolin 61 prec
17	51	36.2	461	2 A31237	alpha-2C-adrenergi
18	51	36.2	574	2 S61234	hypothetical prote
19	51	36.2	705	2 T31157	hypothetical prote
20	51	36.2	965	2 T50670	actin bundling pro
21	51	36.2	1184	2 S50832	atrophin-1 - human
22	50.5	35.8	118	2 S52855	hypothetical prote
23	50.5	35.8	1419	2 T32970	hypothetical prote
24	50	35.5	208	2 T21689	hypothetical prote
25	50	35.5	250	2 T29344	hypothetical prote
26	50	35.5	409	2 S41329	5-enolpyruvylshiki
27	50	35.5	954	1 S68178	mixed-lineage prot
28	49.5	35.1	134	2 D83127	hypothetical prote
29	49.5	35.1	159	2 T03409	MADS box protein -

30	49.5	35.1	168	2 S56755	ZEM2 protein - mai
31	49.5	35.1	209	2 S54209	ZEM1 protein - mai
32	49	34.8	332	2 C83423	conserved hypothet
33	49	34.8	546	1 S13757	RNA helicase DBP2
34	49	34.8	751	2 T48719	hypothetical prote
35	49	34.8	825	2 JC4163	DNA-binding protei
36	49	34.8	1168	1 MWAXIC	myosin heavy chain
37	49	34.8	1275	2 T49362	hypothetical prote
38	48.5	34.4	1733	1 B45344	probable nuclear a
39	48	34.0	103	2 E2709	hypothetical prote
40	48	34.0	138	2 S06706	hypothetical HTLV-
41	48	34.0	156	2 T29730	hypothetical prote
42	48	34.0	212	2 E75334	ribonuclease HII -
43	48	34.0	240	2 T23797	hypothetical prote
44	48	34.0	272	2 T34184	hypothetical prote
45	48	34.0	340	2 A24026	erythromycin resis

ALIGNMENTS

RESULT 1

T36132

hypothetical protein SCE19A.09 - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T36132

R:Seeger, K.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, July 1999

A:Reference number: Z21598

A:Accession: T36132

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-35 <SEE>

A:Cross-references: EMBL:AL096852; PIDN:CAB50990.1; GSPDB:GN00070; SCOEDB:SCE19A.09

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOEDB:SCE19A.09

Query Match

Best Local Similarity 43.3%; Score 61; DB 2; Length 55;

Matches 13; Conservative 50.0%; Pred. No. 0.27;

Mismatches 2; Indels 9; Gaps 1;

QY 2 ESWARGGAPRSGGRC--GRGQVAGP 25

DB 23 DGMGRGGRPRGRGRRPPEGRDERDGP 48

RESULT 2

T09555

fibrillarlin - Arabidopsis thaliana

N:Alternate names: protein L73G19.10

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 15-Oct-1999

C:Accession: T09555

R:Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Sch

submitted to the Protein Sequence Database, May 1999

A:Reference number: Z16098

A:Accession: T09555

A:Molecule type: DNA

A:Residues: 1-320 <AAA>

A:Cross-references: EMBL:AL050400; GSPDB:GN00062; ATSP:L73G19.10

A:Experimental source: cultivar Columbia; BAC clone L73G19

C:Genetics:

A:Gene: ATSP:L73G19.10

A:Map position: 4

A:Introns: 7/1; 120/3; 152/3; 178/3; 237/3; 260/3

C:Superfamily: human fibrillarlin

C:Keywords: nucleus

Query Match

41.8%; Score 59; DB 2; Length 320;

S.: Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999

A::Title: Genome sequence of the radioreistant bacterium Deinococcus radiodurans R1.

A::Reference number: A75250; MUID:20036896

A::Accession: B75450

A::Status: preliminary

A::Molecule type: DNA

A::Residues: 1-210 <WHI>

A::Cross-references: GB:AE001952; GB:AE000513; NID:g6458725; PIDN:AAF10586.1; PID:g645

A::Experimental source: strain R1

C::Genetics:

A::Gene: DR1010

A::Map position: 1

Query Match 39.7%; Score 56; DB 2; Length 210;
Best Local Similarity 53.8%; Pred. No. 3.5;
Matches 14; Conservative 0; Mismatches 8; Indels 4; Gaps 1;

QY 3 SW----ARGGAPRSGGRRCRGQVAG 24
||| ||| ||| |||
Db 135 SWLGRGGRGRRRRGRRACRGORAG 160

RESULT 6

C72662 hypothetical protein APE0723 - Aeropyrum pernix (strain K1)

C::Species: Aeropyrum pernix

C::Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999

C::Accession: C72662

R::Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Tanaka, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.

A::Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero

A::Reference number: A7450; MUID:99310339

A::Accession: C72662

A::Status: preliminary

A::Molecule type: DNA

A::Residues: 1-152 <RAW>

A::Cross-references: DDBJ:AP000060; NID:g5104188; PIDN:BAA79699.1; PID:d1043485; PID:g

A::Experimental source: strain K1

C::Genetics:

A::Gene: APE0723

Query Match 38.3%; Score 54; DB 2; Length 152;
Best Local Similarity 47.6%; Pred. No. 4.8;
Matches 10; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 4 WARGAPRSGGRRCRGQVAG 24
||| | | | | |
Db 49 WERGDGPGGGSRSDGRGPVG 69

RESULT 7

S78110 thi protein - Rhizobium meliloti (fragment)

C::Species: Rhizobium meliloti

C::Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 08-Oct-1999

C::Accession: S78110; S39961

R::Becker, A.

submitted to the EMBL Data Library, December 1994

A::Reference number: S78110

A::Accession: S78110

A::Molecule type: DNA

A::Residues: 1-168 <BEC>

A::Cross-references: EMBL:Z22636; NID:g605653; PIDN:CAA80350.1; PID:g605656

A::Note: This is a revision to the sequence in reference S39955

R::Becker, A.; Kleickmann, A.; Keller, M.; Arnold, W.; Puehler, A.

Mol. Gen. Genet. 241, 367-379, 1993

A::Title: Identification and analysis of the Rhizobium meliloti exoAMONP genes involve

A::Reference number: S39955; MUID:94067019

A::Accession: S39961

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 'D',160-161,'MVATSGDALLRPDAVAALIEELLPLALVVTPLNLAALMTGTAAGDEAMAR','QAEATMR
A;Cross-references: EMBL:222636
A;Note: this sequence has been revised in reference S78110
C;Genetics:
A;Gene: thi
C;Function:
A;Description: involved in thiamine biosynthesis

Query Match 38.3%; Score 54; DB 2; Length 168;
Best Local Similarity 44.0%; Pred. No. 5.2;
Matches 11; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

QY 1 GESWARGAPRSGRRRCRGQVAGP 25
Db 52 GGDGAPGSDHADGRTCGSGQGRAP 76

RESULT 8
A57717
transcription factor EC2 - human
C;Species: Homo sapiens (man)
C;Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 05-Nov-1999
C;Accession: A57717
R;Quaternary, E.E.; Hida, H.; Blau, M.A.; Quaternary, T.
Proc. Natl. Acad. Sci. U.S.A. 91, 7066-7070, 1994
A;Title: Cloning and characterization of a basic helix-loop-helix protein expressed in e
A;Reference number: A57717; MUID:94316638
A;Accession: A57717
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-198 <QUE>
A;Cross-references: GB:008336; NID:g488286; PIDN:AAA19969.1; PID:g514278
C;Genetics:
A;Gene: GDB:TCF15; EC2
A;Cross-references: GDB:387576
A;Map position: 20q12-20q13.1
C;Keywords: transcription factor

Query Match 38.3%; Score 54; DB 2; Length 198;
Best Local Similarity 52.0%; Pred. No. 5.9;
Matches 13; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 GESWARGAPRSGRRRCRGQVAGP 25
Db 43 GPEAARRGPGGGRAGGGGAGP 67

RESULT 9
S14984
glycine-rich protein (clones ua-3 and ua-4) - tomato (fragment)
C;Species: Lycopersicon esculentum (tomato)
C;Date: 13-Jan-1995 #sequence_revision 29-May-1998 #text_change 21-Jul-2000
C;Accession: S14984; S14985
R;Shewalter, A.M.; Zhou, J.; Rumeau, D.; Worst, S.G.; Varner, J.E.
Plant Mol. Biol. 16, 547-565, 1991
A;Title: Tomato extensin and extensin-like cDNAs: structure and expression in response t
A;Reference number: S14970; MUID:91329690
A;Accession: S14984
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-129 <SHO>
A;Cross-references: EMBL:X55695; NID:g19222; PIDN:CAA39224.1; PID:g388258
A;Experimental source: cv. UC82B; haplotype n12; stem tissue; clone ua-3
A;Accession: S14985
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 71-129 <SHW>
A;Cross-references: EMBL:X55696; NID:g19223; PIDN:CAA39225.1; PID:g1345534

A;Experimental source: cv. UC82B; haplotype n12; stem tissue; clone uk-4

Query Match 37.6%; Score 53; DB 2; Length 129;
Best Local Similarity 60.0%; Pred. No. 5.0;
Matches 12; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 GESWARGAPRSGRRRCRG 20
Db 60 GRSGRGGRDGRGGRFGGRG 79

RESULT 10
T33422
hypothetical protein M01G5.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 04-Mar-2000
C;Accession: T33422
R;Geisel, C.; Kramer, J.; Twyman, B.
submitted to the EMBL Data Library, July 1998
A;Description: The sequence of C. elegans cosmid M01G5.
A;Reference number: Z21342
A;Accession: T33422
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-889 <GEI>
A;Cross-references: EMBL:AF078786; PIDN:AAC26941.1; GSPDB:GN00021; CESP:M01G5.1
A;Experimental source: strain Bristol N2; clone M01G5
C;Genetics:
A;Gene: CESP:M01G5.1
A;Map position: 3
A;Introns: 241/3; 356/2; 436/3; 643/2; 674/2; 775/3
C;Superfamily: Caenorhabditis elegans hypothetical protein M01G5.1

Query Match 37.6%; Score 53; DB 2; Length 889;
Best Local Similarity 68.8%; Pred. No. 27;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 6 RGGAPRSGRRRCRG 21
Db 846 RGGRRSGRRGGRGR 861

RESULT 11
B75308
ribonuclease P protein component - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C;Accession: B75308
R;White, O.; Eisen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zaleski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896
A;Accession: B75308
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-166 <WHI>
A;Cross-references: GB:AE002049; GB:AE000513; NID:g6459945; PIDN:AAF11699.1; PID:g645
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR2151
A;Map position: 1

Query Match 37.2%; Score 52.5; DB 2; Length 166;
Best Local Similarity 63.2%; Pred. No. 7.9;
Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 6 RGGAPRSGRRRCRGQVAG 24
Db 1111 1111 1111 1111

Db 119 RGAAGKGGKK-GGQVAG 136

RESULT 12

S07132

hypothetical protein 1 - Rhodospirillum rubrum (fragment)

C:Species: Rhodospirillum rubrum

C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 08-Oct-1999

C:Accession: S07132

R:Falk, G.; Hampe, A.; Walker, J.E.

Biochem. J. 228, 391-407, 1985

A:Title: Nucleotide sequence of the Rhodospirillum rubrum atp operon.

A:Reference number: S07132; MUID:85251588

A:Accession: S07132

A:Molecule type: DNA

A:Residues: 1-886 <FAL>

A:Cross-references: EMBL:X02499; NID:946360; PIDN:CAA26335.1; PID:e10835; PID:g1333809

Query Match 37.2%; Score 52.5; DB 2; Length 886;

Best Local Similarity 44.8%; Pred. No. 31;

Matches 13; Conservative 0; Mismatches 7; Indels 9; Gaps 1;

Qy 6 RGAAPRSGG-----RRCGRGQVAG 25

Db 117 RGRRRRGGRELLGLVVRPAGRGQTGCP 145

RESULT 13

S33690

fibrillarin - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 10-Dec-1999

C:Accession: S33690; T40113

R:Girard, J.P.; Fellu, J.; Caizergues-Ferrer, M.; Lapeyre, B.

Nucleic Acids Res. 21, 1881-1887, 1993

A:Title: Study of multiple fibrillarin mRNAs reveals that 3' end formation in Schizosaccharomyces pombe involves a specific 3' end formation in Schizosaccharomyces pombe

A:Reference number: S33690; MUID:93261822

A:Accession: S33690

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-305 <GIR>

A:Cross-references: EMBL:X69930; NID:g296703; PIDN:CAA49550.1; PID:g296704

R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Taylor, K.; Harris, D.

submitted to the EMBL Data Library, September 1998

A:Reference number: Z21906

A:Accession: T40113

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-305 <WO>

A:Cross-references: EMBL:AL031788; PIDN:CAA21168.1; GSPDB:GN00067; SPDB:SPBC2D10.10C

A:Experimental source: strain 972h-; cosmid c2D10

C:Genetics:

A:Gene: SPBC2D10.10C

A:Map position: 2

C:Superfamily: human fibrillarin

C:Keywords: nucleus

Query Match 36.9%; Score 52; DB 2; Length 305;

Best Local Similarity 50.0%; Pred. No. 15;

Matches 12; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

Qy 1 GESWARGAPRSGRGCGQVAG 24

Db 39 GRCGARGGRGGRGGRGSSG 62

RESULT 14

S25421

nucleolar protein NOP1 - yeast (Saccharomyces cerevisiae)

N:Alternate names: nucleolar protein, 38K; protein D2870; protein YDL014W

C:Species: Saccharomyces cerevisiae

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 20-Jun-2000

C:Accession: S25421; A35038; B35038; S52505; S67546

R:Schimmang, T.; Tollervey, D.; Kern, H.; Frank, R.; Hurt, E.C.

EMBO J. 8, 4015-4024, 1989

A:Title: A yeast nucleolar protein related to mammalian fibrillarin is associated with the nucleolar protein NOP1

A:Reference number: S25421; MUID:90076121

A:Accession: S25421

A:Molecule type: DNA

A:Residues: 1-327 <SCH>

A:Cross-references: EMBL:X51676

R:Henriquez, R.; Blobel, G.; Aris, J.P.

J. Biol. Chem. 265, 2209-2215, 1990

A:Title: Isolation and sequencing of NOP1. A yeast gene encoding a nucleolar protein

A:Reference number: A35038; MUID:90130477

A:Accession: A35038

A:Molecule type: DNA

A:Residues: 1-327 <HEN>

A:Cross-references: EMBL:J05230; NID:g172047; PIDN:AAA34816.1; PID:g172048

A:Accession: B35038

A:Molecule type: protein

A:Residues: 112-142; 157-181; 211-231 <HE2>

R:Andre, B.; Vissers, S.; Urrestarazu, L.

submitted to the EMBL Data Library, February 1995

A:Description: The sequence of a 42 kb segment located on the left arm of chromosome 1

A:Reference number: S52492

A:Accession: S52505

A:Molecule type: DNA

A:Residues: 1-327 <AND>

A:Cross-references: EMBL:Z48432; NID:g683669; PIDN:CAA88345.1; PID:g683683

R:Urrestarazu, L.A.; Andre, B.; Vissers, S.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S67535

A:Accession: S67546

A:Molecule type: DNA

A:Residues: 1-327 <URR>

A:Cross-references: EMBL:Z74062; NID:g1430978; PIDN:CAA98572.1; PID:g1430979; MIPS:YD

A:Experimental source: strain S288C

C:Genetics:

A:Gene: SGD:NOP1

A:Cross-references: SGD:S0002172; MIPS:YDL014W

A:Map position: 4L

C:Superfamily: human fibrillarin

C:Keywords: nucleus

Query Match 36.9%; Score 52; DB 2; Length 327;

Best Local Similarity 54.2%; Pred. No. 16;

Matches 13; Conservative 1; Mismatches 8; Indels 2; Gaps 1;

Qy 1 GESWARGAPRSGRGCGQVAG 24

Db 57 GRGSRGGA--RGGSRGGRGGAAG 78

RESULT 15

C70743

hypothetical protein Rv0480c - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

C:Accession: C70743

R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon

Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987

A:Accession: C70743

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-340 <COL>

A:Cross-references: GB:Z77162; GB:AL123456; NID:g3261606; PIDN:CAB00941.1; PID:e25514

A:Experimental source: strain H37RV

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 17:06:14 ; Search time 80.04 seconds
(without alignments)
10.087 Million cell updates/sec

Title: US-09-518-931-4_COPY_L42_166
Perfect score: 141
Sequence: 1 GESWARGGAPRSGRRGRGVAGP 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	52.5	37.2	166	1 RNPA_DEIRA	Qrsh3 Deinococcus
2	52	36.9	305	1 FBRL_SCHPO	P35551 schizosacch
3	52	36.9	327	1 FBRL_YEAST	P15646 saccharomyc
4	52	36.9	340	1 Y480_MYCTU	Q11146 mycobacteri
5	51	36.2	71	1 CBA_CARP1	P38578 carnobacter
6	51	36.2	461	1 A2AK_HUMAN	P18825 homo sapien
7	50.5	35.8	594	1 APAL_MOUSE	P97459 mus musculu
8	50	35.5	297	1 FBRL_LEIMA	P35549 leishmania
9	50	35.5	954	1 M3KA_HUMAN	Q02779 homo sapien
10	49	34.8	409	1 AROA_BURPS	P39915 burkholderi
11	49	34.8	410	1 AA2A_MOUSE	Q06613 mus musculu
12	49	34.8	546	1 DBP2_YEAST	P24783 saccharomyc
13	49	34.8	825	1 SE5_RAT	Q63003 rattus norv
14	49	34.8	1168	1 MYSC_ACACA	P10569 acanthamoeb
15	48.5	34.4	508	1 Y202_HUMAN	Q92599 homo sapien
16	48.5	34.4	1733	1 VNVA_PRYKA	P33485 pseudorabie
17	48	34.0	208	1 Y4IN_RHLSN	P55497 rhizobium s
18	48	34.0	263	1 RNH2_BACHD	Q92980 bacillus ha
19	48	34.0	272	1 RS2_CAEEL	P51403 caenorhabdi
20	48	34.0	340	1 RMAA_ARTS3	P09891 arthrobacte
21	48	34.0	342	1 ROAL_SCHAM	P21522 schistocerc
22	48	34.0	427	1 GBP2_YEAST	P25555 saccharomyc
23	48	34.0	485	1 ST13_SCHPO	Q09181 schizosacch
24	48	34.0	543	1 YP91_MYCTU	Q50630 mycobacteri
25	48	34.0	942	1 KDGD_HUMAN	P29284 homo sapien
26	47.5	33.7	412	1 AA2A_HUMAN	P29274 homo sapien
27	47.5	33.7	461	1 D15K_MOUSE	Q61466 mus musculu
28	47.5	33.7	470	1 EGR2_MOUSE	P08152 mus musculu
29	47.5	33.7	470	1 EGR2_RAT	P31774 rattus norv
30	47	33.3	558	1 YJ83_MYCTU	Q10873 mycobacteri
31	47	33.3	650	1 NUCLE_XENLA	P20397 xenopus lae
32	46.5	33.0	388	1 PGK_MYCAV	Q32848 mycobacteri
33	46	32.6	227	1 EMBP_RAT	Q63189 rattus norv

34	46	32.6	285	1	RK2_SPIOI	P06509 spinacia ol
35	46	32.6	567	1	Y095_SYNY3	Q55884 synechocyst
36	46	32.6	676	1	CIQI_HUMAN	P51787 homo sapien
37	46	32.6	676	1	ICP0_HSVBK	P29836 bovine herp
38	46	32.6	706	1	NUCL_HUMAN	P19338 homo sapien
39	46	32.6	706	1	NUCL_MOUSE	P09405 mus musculu
40	46	32.6	712	1	NUCL_RAT	P13383 rattus norv
41	45.5	32.3	247	1	NP24_LYCES	P12670 lycopersico
42	45.5	32.3	314	1	CHIB_VITVI	P51613 vitis vinif
43	45.5	32.3	403	1	H136_ARATH	O82660 arabidopsis
44	45.5	32.3	421	1	NUF2_RHIME	P36913 rhizobium m
45	45.5	32.3	713	1	NUCL_MESAU	P08199 mesocricetu

ALIGNMENTS

```
RESULT 1
RNPA_DEIRA
ID RNPA_DEIRA STANDARD; PRT; 166 AA.
AC Q9RSH3:
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE RIBONUCLEASE P (Protein component (EC 3.1.26.5) (Protein C5) (RNase P)).
GN RNPA OR DR2151.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Hart D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Oin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1."
RL Science 286:1571-1577(1999).
CC -!- FUNCTION: RIBONUCLEASE P GENERATES MATURE TRNA MOLECULES BY
CC CLEAVING THEIR 5' ENDS. IT CAN CLEAVE ALSO THE 4.5S RNA
CC (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ENDONUCLEOLYTIC CLEAVAGE OF RNA, REMOVING
CC 5'-EXTRA-NUCLEOTIDE FROM TRNA PRECURSOR.
CC -!- MISCELLANEOUS: RNASE P CONSISTS OF A RNA MOIETY (M1, RNBP) AND THE
CC PROTEIN COMPONENT. BOTH ARE NECESSARY FOR FULL ENZYMATIC ACTIVITY.
CC HOWEVER, IT IS THE RNA THAT CARRIES THE CATALYTIC SITE.
CC -!- SIMILARITY: BELONGS TO THE RNPA FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE002049; AAF11699.1; -.
DR TIGR; DR2151; -.
DR INTERPRO; IPR000100; -.
DR PFAM; PF00825; Ribonuclease_P; 1.
DR PROSITE; PS00648; RIBONUCLEASE_P; 1.
DR Hydrolase; Nuclease; Endonuclease; tRNA processing.
SQ SEQUENCE 166 AA; 17772 MW; 4DD4055F8F2E5042 CRC64;
```

Query Match 37.2%; Score 52.5; DB 1; Length 166;
Best Local Similarity 63.2%; Pred. No. 2.9;
Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

```

QY      6  RGGAPRSGRRGRCRGQVAG 24
DB      119  RGCAGKGGKK-GGGQVAG 136

RESULT 2
FBRL SCHPO
ID FBRL SCHPO STANDARD; PRT; 305 AA.
AC P3551;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE FIBRILLARIN.
GN FIB OR SPBC2D10.10C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;
OC Schizosaccharomycetaceae; Schizosaccharomycetes.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=93261822; PubMed=8493104;
RA Girard J.-P., Fellu J., Caizerques-Ferrer M., Lapeyre B.;
RT "Study of multiple fibrillar mRNAs reveals that 3' end formation in
RL Schizosaccharomycetes pombe is sensitive to cold shock.";
RL Nucleic Acids Res. 21:1881-1887(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Wood V., Rajandream M.A., Barrell B.G., Taylor K., Harris D.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: FIBRILLARIN IS A COMPONENT OF A NUCLEOLAR SMALL NUCLEAR
CC RIBONUCLEOPROTEIN PARTICLE THOUGHT TO PARTICIPATE IN THE FIRST
CC STEP IN PROCESSING PRERIBOSOMAL RNA.
CC -!- SUBCELLULAR LOCATION: NUCLEAR; FIBRILLAR REGION OF THE NUCLEOLUS.
CC -!- PTM: BY HOMOLOGY TO OTHER FIBRILLARINS, SOME OR ALL OF THE
CC N-TERMINAL ARGININES ARE N,N-DIMETHYLATED (DMA).
CC -!- SIMILARITY: BELONGS TO THE FIBRILLARIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X69930; CAA49550.1; -
DR AL031788; CAA21168.1; -
DR F01269; S33690.
DR INTERPRO; IPR000692; -
DR PFAM; PF01269; Fibrillarin; 1.
DR PRINTS; PR00052; FIBRILLARIN.
DR PROSITE; PS00566; FIBRILLARIN; 1.
KW Nuclear protein; Methylation; Ribonucleoprotein; rRNA processing;
KW RNA-binding.
FT DOMAIN 6 70 GLY/DMA-RICH.
FT DOMAIN 173 180 RNA-BINDING (RNP2) (POTENTIAL).
FT SEQUENCE 305 AA; 32040 MW; 5ABE8B37EC33331 CRC64;

Query Match 36.9%; Score 52; DB 1; Length 305;
Best Local Similarity 50.0%; Pred. No. 5.6;
Matches 12; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY      1  GESWARGAPRSGRRGRCRGQVAG 24
DB      39  GRGARGRGRRGARGRGSSG 62

RESULT 3
FBRL YEAST
ID FBRL YEAST STANDARD; PRT; 327 AA.
AC P15646; P89B90;

```

```

DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE FIBRILLARIN (NUCLEOLAR PROTEIN 1).
GN NOPI OR YDL014W OR D2870.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetaceae; Saccharomycetes.
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC STRAIN=BJ2168;
RX MEDLINE=90130477; PubMed=2298745;
RA Henriquez R., Blobel G., Aris J.P.;
RT "Isolation and sequencing of NOPI. A yeast gene encoding a nucleolar
RL protein homologous to a human autoimmune antigen.";
RL J. Biol. Chem. 265:2209-2215(1990).
RN [2]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RX MEDLINE=90076121; PubMed=2686980;
RA Schimang T., Tollervey D., Kern H., Frank R., Hurt E.C.;
RT "A yeast nucleolar protein related to mammalian fibrillarin is
RL associated with small nucleolar RNA and is essential for viability.";
RL EMBO J. 8:4015-4024(1989).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RA Andre B., Vissers S., Urrestazu L.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RP FUNCTION.
RX MEDLINE=91160511; PubMed=1825809;
RA Tollervey D., Lehtonen H., Carmo-Fonseca M., Hurt E.C.;
RT "The small nucleolar RNP protein NOPI (fibrillarin) is required for
RL pre-rRNA processing in yeast.";
RL EMBO J. 10:573-583(1991).
CC -!- FUNCTION: FIBRILLARIN IS A COMPONENT OF A NUCLEOLAR SMALL NUCLEAR
CC RIBONUCLEOPROTEIN PARTICLE THOUGHT TO PARTICIPATE IN THE FIRST
CC STEP IN PROCESSING PRERIBOSOMAL RNA.
CC -!- SUBCELLULAR LOCATION: NUCLEAR; FIBRILLAR REGION OF THE NUCLEOLUS.
CC -!- PTM: BY HOMOLOGY TO OTHER FIBRILLARINS, SOME OR ALL OF THE
CC N-TERMINAL ARGININES ARE N,N-DIMETHYLATED (DMA).
CC -!- SIMILARITY: BELONGS TO THE FIBRILLARIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J05230; AAA34816.1; -
DR EMBL; 248432; CAA88345.1; -
DR EMBL; 274061; CAA98571.1; -
DR EMBL; 274062; CAA98572.1; -
DR F01269; S35038.
DR F01269; S25421; S25421.
DR SGD; S0002172; NOPI.
DR INTERPRO; IPR000692; -
DR PFAM; PF01269; Fibrillarin; 1.
DR PRINTS; PR00052; FIBRILLARIN.
DR PROSITE; PS00566; FIBRILLARIN; 1.
KW Nuclear protein; Methylation; Ribonucleoprotein; rRNA processing;
KW RNA-binding.
FT DOMAIN 6 83 GLY/DMA-RICH.
FT DOMAIN 22 42 RNA-BINDING RGG-BOX (BY SIMILARITY).
FT DOMAIN 58 75 RNA-BINDING RGG-BOX (BY SIMILARITY).
FT DOMAIN 192 199 RNA-BINDING (RNP2) (POTENTIAL).
FT DOMAIN 281 313 HELICAL (POTENTIAL).
FT SEQUENCE 327 AA; 34465 MW; 56A8B958A7B6066E CRC64;

Query Match 36.9%; Score 52; DB 1; Length 327;

```


RN AC P97459;
 RP TISSUE=FROM N.A.
 RC MEDLINE=98041882; PubMed=9371698;
 RX Schack S., Davedjian J.C., Cayla C., Sender Y., Paris H.;
 RA "Molecular cloning, sequencing and functional study of the promoter
 RT region of the human alpha2c4-adrenergic receptor gene.";
 RL Biochem. J. 328:431-438(1997).
 CC -!- FUNCTION: ALPHA-2 ADRENERGIC RECEPTORS MEDIATE THE CATECHOLAMINE-
 CC INDUCED INHIBITION OF ADENYLATE CYCLASE THROUGH THE ACTION OF G
 CC PROTEINS.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J03853; AAA35513.1; -;
 DR EMBL; U72648; AAC78723.1; -;
 DR PIR; A31237; A31237.
 DR HSSP; P29274; 1MMH.
 DR GCRDB; GCR_0044; -;
 DR GCRDB; GCR_1903; -;
 DR MIM; 104250; -;
 DR INTERPRO; IPR000276; -;
 DR INTERPRO; IPR000735; -;
 DR INTERPRO; IPR002233; -;
 DR PFAM; PF00001; 7tm1.1;
 DR PRINTS; PR00237; GPCRHHODPSN.
 DR PRINTS; PR00560; ADRENRCAC2R.
 DR PRINTS; PR01103; ADRENERGICR.
 DR PROSITE; PS00237; G-PROTEIN_RECP_F1_1; 1.
 DR PROSITE; PS0262; G-PROTEIN_RECP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Multigene family; Phosphorylation.
 FT DOMAIN 1 51
 FT TRANSMEM 52 76
 FT DOMAIN 77 88
 FT TRANSMEM 89 114
 FT DOMAIN 115 124
 FT TRANSMEM 125 147
 FT DOMAIN 148 168
 FT TRANSMEM 169 191
 FT DOMAIN 192 207
 FT TRANSMEM 208 231
 FT DOMAIN 232 333
 FT TRANSMEM 383 406
 FT DOMAIN 407 419
 FT TRANSMEM 420 440
 FT DOMAIN 441 461
 FT CARBOHYD 19 19
 FT CARBOHYD 33 33
 FT DISULFID 124 202
 SQ SEQUENCE 461 AA; 49513 MW; 7BA861A4AD1200AE CRC64;
 Query Match 36.2%; Score 51; DB 1; Length 461;
 Best Local Similarity 46.9%; Pred. No. 11;
 Matches 15; Conservative 0; Mismatches 5; Indels 12; Gaps 2;
 QY 6 RCGAP-----RSGGRRRCGR-----GQVAGP 25
 DB 293 RCGAGPLRRGRRRRRAGAGGAGGAGGQAGP 324
 RESULT 7
 NPAL_MOUSE STANDARD; PRT; 594 AA.
 ID NPAL_MOUSE

AC P97459;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DE NEURONAL PAS DOMAIN PROTEIN 1 (NEURONAL PAS1).
 GN NPAS1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RX MEDLINE=97163088; PubMed=9012850;
 RA Zhou Y.-D., Barnard M., Tian H., Li X., Ring H.Z., Francke U.,
 RT Shelton J., Richardson J., Russell D.W., McKnight S.L.;
 RT "Molecular characterization of two mammalian bHLH-PAS domain proteins
 RT selectively expressed in the central nervous system.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:713-718(1997).
 CC -!- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
 CC BHLH PROTEIN.
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
 CC -!- TISSUE SPECIFICITY: IN BRAIN, EXCLUSIVELY NEURONAL. ALSO FOUND IN
 CC SPINAL CORD.
 CC -!- DEVELOPMENTAL STAGE: FIRST DETECTED BETWEEN EMBRYONIC DAY 15 AND
 CC DAY 16.
 CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
 CC TRANSCRIPTION FACTORS.
 CC -!- SIMILARITY: CONTAINS 1 PAS (PER-ARNT-SIM) DIMERIZATION DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U7967; AAB47247.1; -;
 DR MGD; MGI:109205; NPAS1.
 DR INTERPRO; IPR000014; -;
 DR INTERPRO; IPR003015; -;
 DR PFAM; PF00989; PAS; 1.
 DR PROSITE; PS00038; HELIX_LOOP_HELIX; 1.
 KW Repeat; DNA-binding; Nuclear protein; Transcription regulation.
 FT DNA_BIND 46 58
 FT DOMAIN 59 99
 FT REPEAT 137 203
 FT REPEAT 294 360
 FT DOMAIN 366 409
 FT DOMAIN 212 223
 FT POLY-SER.
 SQ SEQUENCE 594 AA; 63736 MW; D6477696DF69A4B3 CRC64;
 Query Match 35.8%; Score 50.5; DB 1; Length 594;
 Best Local Similarity 61.1%; Pred. No. 16;
 Matches 11; Conservative 2; Mismatches 2; Indels 3; Gaps 1;
 QY 10 PRSGGR--RCGRGOVAG 24
 DB 6 PRSGGRGEVKCGGRGAG 23
 RESULT 8
 FBRL_LEIMA STANDARD; PRT; 297 AA.
 ID FBRL_LEIMA
 AC P35549;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE FIBRILLARIN.
 OS Leishmania major.
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 RN [1]

RA MEDLINE=88016163; PubMed=3477803;
RT Jung G., Korn E.D., Hammer J.A. III;
RT "The heavy chain of Acanthamoeba myosin IB is a fusion of myosin-like
RT and non-myosin-like sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 84:6720-6724(1987).
RN [2]
RP PARTIAL SEQUENCE FROM N.A.
RX MEDLINE=86259656; PubMed=3014500;
RA Hammer J.A. III, Jung G., Korn E.D.;
RT "Genetic evidence that Acanthamoeba myosin I is a true myosin.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:4655-4659(1986).
RN [3]
RP PHOSPHORYLATION SITE.
RX MEDLINE=90037074; PubMed=2530230;
RA Brzeska H., Lynch T.J., Martin B., Korn E.D.;
RT "The localization and sequence of the phosphorylation sites of
RT Acanthamoeba myosins I. An improved method for locating the
RT phosphorylated amino acid.";
RL J. Biol. Chem. 264:19340-19348(1989).
CC -!- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE
CC -!- SUBUNIT: MYOSIN I HEAVY CHAIN IS SINGLE-HEADED. DIMER OF A HEAVY
CC AND A LIGHT CHAIN. INABILITY TO SELF-ASSEMBLE INTO FILAMENTS;
CC MYOSINS I CAN THEREFORE MOVE ACTIN RELATIVE TO MEMBRANES;
CC VERSA. TH.2 AND SH3 BIND TIGHTLY TO F-ACTIN; THIS TOGETHER WITH
CC THE NUCLEOTIDE-SENSITIVE SITE IN THE HEAD, ALLOWS SINGLE MOLECULES
CC OF MYOSIN I TO CROSS-LINK ACTIN FILAMENTS.
CC -!- MISCELLANEOUS: THIS ORGANISM EXPRESSES AT LEAST THREE ISOFORMS OF
CC MYOSIN I HEAVY-CHAIN, ENCODED BY GENES MIA, MIB, AND MIC.
CC -!- SIMILARITY: THE GLOBULAR HEAD DOMAIN IS SIMILAR WITH THE GLOBULAR
CC HEAD SEQUENCE OF MUSCLE MYOSINS. BY CONTRAST, THE COILED-COIL ROD-
CC LIKE REGION FOUND IN MANY MYOSIN HEAVY CHAINS IS REPLACED BY A
CC CARBOXYL-TERMINAL DOMAIN CONTAINING A REGION RICH IN GLYCINE,
CC ALANINE, AND PROLINE.
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -!- CAUTION: WAS ORIGINALLY THOUGHT TO BE MYOSIN IB.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J02974; AAA27707.1; -
DR PIR; A33891; MWXIC.
DR HSSP; P08799; 1MND.
DR INTERPRO; IPR001452; -
DR PFAM; PF00018; SH3; 1.
DR PRINTS; PR00063; myosin.head; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR PRINTS; PR00452; SH3DOMAIN.
DR PROSITE; PS50002; SH3; 1.
KW Myosin; ATP-binding; Phosphorylation; Multigene family; SH3 domain.
FT DOMAIN 1 670 GLOBULAR HEAD-LIKE DOMAIN.
FT DOMAIN 671 1168 NON ALPHA-HELICAL, C-TERMINAL DOMAIN.
FT NP_BIND 101 108 ATP (POTENTIAL).
FT DOMAIN 671 922 TAIL HOMOLGY REGION 1 (TH.1).
FT DOMAIN 923 978 GLY/PRO/ALA-RICH (TH.2).
FT DOMAIN 1034 1168 GLY/PRO/ALA-RICH (TH.2).
FT DOMAIN 976 1035 SH3.
FT MOD_RES 311 311 PHOSPHORYLATION.
SQ SEQUENCE 1168 AA; 127309 MW; D07084B373A37A32 CRC64;

Query Match 34.8%; Score 49; DB 1; Length 1168;
Best Local Similarity 65.0%; Pred. No. 44;
Matches 13; Conservative 0; Mismatches 3; Indels 4; Gaps 2;

QY 6 RGGAPRSGRGRCRGVAGP 25

Db 1126 RGGAPPPGGMGR-GRG---GP 1141
||||| ||| ||| ||
RESULT 15
Y202_HUMAN
ID Y202_HUMAN STANDARD; PRT; 508 AA.
AC Q92599;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE SEPTIN-LIKE PROTEIN KIAA0202 (FRAGMENT).
GN KIAA0202.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BONE MARROW;
RX MEDLINE=97191544; PubMed=9039502;
RA Nagase T., Seki N., Ishikawa K.-I., Ohira M., Kawarabayashi Y.,
RA Ohara O., Tanaka A., Kotani H., Miyajima N., Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. VI.
RT The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by
RT analysis of cDNA clones from cell line KG-1 and brain.";
RL DNA RES. 3:321-329(1996).
CC -!- SIMILARITY: BELONGS TO THE SEPTIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D86957; BAA13193.1; -
DR INTERPRO; IPR000038; -
DR PFAM; PF00735; GTP_CDC; 1.
KW Hypothetical protein; GTP-binding; Coiled coil.
FT NON_TER 1 1
FT NP_BIND 130 137 GTP (POTENTIAL).
FT DOMAIN 339 492 COILED COIL (POTENTIAL).
SQ SEQUENCE 508 AA; 58011 MW; A15AB2F91C79FA68 CRC64;

Query Match 34.4%; Score 48.5; DB 1; Length 508;
Best Local Similarity 61.9%; Pred. No. 25;
Matches 13; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 1 GFSWARGGAPRSGRRCGRGQ 21
| ||| | ||| |||
Db 3 GSGCARGRAGR-GGRSRGRGQ 22

Search completed: January 30, 2001, 17:06:15
Job time: 1236 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:55:04 ; Search time 254.27 Seconds
(without alignments)
11.524 Million cell updates/sec

Title: US-09-518-931-4_COPY_142_166
Perfect score: 141
Sequence: 1 GESWARGAPRRGRRGQVAGP 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_15:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	61	43.3	55	2 Q9S2H2	Q9s2h2 streptomyc
2	59	41.8	320	10 Q9SZ21	Q9sz21 arabidopsis
3	57.5	40.8	199	2 Q9X8U3	Q9x8u3 streptomyc
4	57	40.4	562	4 Q9Y583	Q9y583 homo sapien
5	56	39.7	210	2 Q9RVL7	Q9rvl7 deinococcus
6	54	38.3	152	1 Q9YE47	Q9ye47 aeropyrum p
7	54	38.3	198	4 Q12870	Q12870 homo sapien
8	54	38.3	199	4 Q9NQ01	Q9nq01 homo sapien
9	53	37.6	129	10 Q02090	Q02090 lycopersico
10	53	37.6	342	5 Q61580	Q61580 drosophila
11	53	37.6	889	5 Q76686	Q76686 caenorhabdi
12	53	37.6	1912	12 Q9J5C3	Q9j5c3 fowlpox vir
13	53	37.6	2499	5 Q9N9N0	Q9n9n0 leishmania
14	52.5	37.2	104	12 Q9IBX3	Q9ibx3 turkey herp
15	52.5	37.2	124	12 Q85587	Q85587 gallid herp
16	52.5	37.2	124	12 Q9IBS0	Q9ibs0 turkey herp
17	52.5	37.2	886	2 Q9LCV5	Q9lcv5 rhodospiril
18	51.5	36.5	254	10 Q9LI08	Q9li08 oryza sativ
19	51	36.2	71	2 Q9REZ1	Q9rez1 carnobacter

Q9njb6 trypanosoma
Q9lda8 oryza sativ
Q65563 bovine herp
Q85866 sphingomona
Q9sqh4 lilium long
Q9ul43 leishmania
Q9ul09 leishmania
Q43959 azorhizobiu
Q91594 xenopus lae
P70000 xenopus lae
Q9wqg5 bufo arenar
Q9sm06 arabidopsis
Q75040 homo sapien
Q45092 caenorhabdi
Q46029 chironomus
Q62213 caenorhabdi
Q19667 caenorhabdi
Q9vhp0 drosophila
Q9wty8 rattus norv
Q15061 homo sapien
Q41828 zea mays (m
Q41893 zea mays (m
Q41892 zea mays (m
Q41891 zea mays (m
Q9ul12 leishmania
Q9qyf4 mus musculu

ALIGNMENTS

RESULT 1

Q9S2H2 ID Q9S2H2 PRELIMINARY; PRT; 55 AA.

AC Q9S2H2;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
DE HYPOTHEICAL 6.0 KDA PROTEIN.

GN SCE19A.09.

OS Streptomyces coelicolor.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

OX NCBI_TaxID=1902;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-A3(2);

RA Seeger K., Harris D.;

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-A3(2);

RA James K.D., Parkhill J., Barrell B.G., Rajandream M.A.;

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN-A3(2);

RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,

RA Kinashi H., Hopwood D.A.;

RT "A set of ordered cosmids and a detailed genetic and physical map for

the 8 Mb Streptomyces coelicolor A3(2) chromosome.";

RL Mol. Microbiol. 21:77-96(1996).

DR EMBL: AL096852; CAB50990.1; -;

KW Hypothetical protein.

SQ SEQUENCE 55 AA; 5971 MW; C5C2BECDB0A253F9 CRC64;

Query Match 43.3%; Score 61; DB 2; Length 55;

Best Local Similarity 50.0%; Pred. No. 0.2;

Matches 13; Conservative 2; Mismatches 9; Indels 2; Gaps 1;

QY 2 ESWARGGAPRRGRRGQVAGP 25

: | | | | | | | | | | | | | | | | | | | |

ID Q02090 PRELIMINARY; PRT; 129 AA.
AC Q02090;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE GLYCINE-RICH PROTEIN (FRAGMENT).
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. UC82B; TISSUP=STEM;
RX MEDLINE=91329690; PubMed=1714316;
RA Shewalter A.M., Zhou J., Rumeau D., Worst S.G., Varner J.E.;
RT "Tomato extensin and extensin-like cDNAs: structure and expression in
RT response to wounding.";
RL Plant Mol. Biol. 16:547-565(1991).
DR EMBL; X55695; CAA39224.1; -;
DR PIR; S14984; S14984.
DR MENDEL; 15241; Lyces; 2406; 15241.
DR INTERPRO; IPR000504; -;
DR PROSITE; PS00030; RNP_1; UNKNOWN_1.
FT NON_TER 1 129
FT NON_TER 1 129
SQ SEQUENCE 129 AA; 12495 MW; A508A004129A9E9B CRC64;
Query Match 37.6%; Score 53; DB 10; Length 129;
Best Local Similarity 60.0%; Pred. No. 5.2;
Matches 12; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 GESWARGAPRSGGRCGRG 20
Db 60 GRSGRGGRGDSGGRGGRG 79
RESULT 10
ID O61580 PRELIMINARY; PRT; 342 AA.
AC O61580;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE Y BOX PROTEIN.
GN YB-1.
OS Drosophila silvestris (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=47010;
RN [1]
RP SEQUENCE FROM N.A.
RA Davis T., Kuribara J., Yamamoto D.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.
DR EMBL; AF052296; AAC06034.1; -;
DR HSSP; PL5277; 1MJC.
DR FLYBASE; FBgn0024359; Dsill\YB-1.
DR INTERPRO; IPR002059; -;
DR PFAM; PF00313; CSD; 1.
DR PRINTS; PR00050; COLDSHOCK.
DR PROSITE; PS00352; COLD_SHOCK; 1.
DR PRODOM; PD000621; -; 1.
KW Transcription regulation; DNA-binding.
SQ SEQUENCE 342 AA; 37051 MW; EAA91925E6053F89 CRC64;
Query Match 37.6%; Score 53; DB 5; Length 342;
Best Local Similarity 60.0%; Pred. No. 13;
Matches 12; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 6 RCGAPRSGRCRCQVACP 25
Db 202 RCGAPRGGGGPGGPPGAGP 221
RESULT 11
O76686 PRELIMINARY; PRT; 889 AA.
ID O76686;
AC O76686;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE MOIG5.1 PROTEIN.
GN MOIG5.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Consey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Geisel C., Kramer J., Twyman B.;
RT "The sequence of C. elegans cosmid MOIG5.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF078786; AAC26941.1; -;
SQ SEQUENCE 889 AA; 102060 MW; 3A0B803B3133D39B CRC64;
Query Match 37.6%; Score 53; DB 5; Length 889;
Best Local Similarity 68.8%; Pred. No. 32;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 6 RCGAPRSGRCRCRCQ 21
Db 846 RCGRGRSGRGGGRGR 861
RESULT 12
Q9J5C3 PRELIMINARY; PRT; 1912 AA.
ID Q9J5C3
AC Q9J5C3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE ORF FPV097 VARIOLA B22R GENE FAMILY PROTEIN.
GN FPV097.
OS Fowlpox virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Avipoxvirus.
OX NCBI_TaxID=10261;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=20193820; PubMed=10729156;
RA Afonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT "The genome of fowlpox virus."
RL J. Virol. 74:3815-3831(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Afonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF198100; AAF4441.1; -;
SQ SEQUENCE 1912 AA; 215979 MW; 9C4C709F671826FE CRC64;

Query Match 37.6%; Score 53; DB 12; Length 1912;
Best Local Similarity 57.9%; Pred. No. 66;
Matches 11; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 GESWARGAPRSGGRCRCR 19
I ||| I ||| I ||
Db 874 GRGGGRRGGRRGRRGGR 892

RESULT 13

Q9N9N0 PRELIMINARY; PRT; 2499 AA.
AC Q9N9N0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE HYPOTHETICAL 263.6 KDA PROTEIN.
GN L3302.03.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FRIEDLIN;
RA Hilbert H., Wedler E., Duesterhoeft A., Ivens A.C.,
RA Quail M., Rajandream M.A., Barrell B.G.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FRIEDLIN;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RT "A physical map of the Leishmania major Friedlin genome."
RL Genome Res. 8:135-145(1998).
DR EMBL: AL359781; CAB95306.1; -;
KW Hypothetical protein.
SQ SEQUENCE 2499 AA; 263616 MW; 81141DEF39126811 CRC64;

Query Match 37.6%; Score 53; DB 5; Length 2499;
Best Local Similarity 45.5%; Pred. No. 85;
Matches 10; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 3 SWARGAPRSGRCRCQVAG 24
I ||| I ||| I ||| I ||
Db 1188 SWTVDNPPSGGRCMCGSGVDG 1209

RESULT 14

Q9IBX3 PRELIMINARY; PRT; 104 AA.
AC Q9IBX3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE R-LORE13.
GN R-LORE13.
OS Turkey herpesvirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae.
OX NCBI_TaxID=10390;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GA;
RX MEDLINE=92237304; PubMed=1315048;
RA Jones D., Lee L., Liu J.L., Kung H.J., Tillotson J.K.;
RT "Marek disease virus encodes a basic-leucine zipper gene resembling
RT the fos/jun oncogenes that is highly expressed in lymphoblastoid
RT tumors."
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=GA;
RA Lee L.F., Wu P., Sui D., Ren D., Kung H.J.;
RT "The Complete UL Sequence of Serotype I Marek's Disease Virus."
RL Proc. Natl. Acad. Sci. U.S.A. 0:0-0(2000).
DR EMBL: AF147806; AAF66809.1; -;
SQ SEQUENCE 104 AA; 11899 MW; 2F7FC36E6FD927E0 CRC64;

Query Match 37.2%; Score 52.5; DB 12; Length 104;
Best Local Similarity 68.8%; Pred. No. 4.9;
Matches 11; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 4 WARGGAPRSGGRCRCR 19
I ||| I ||| I ||
Db 58 WKRGGPRGRRR-GR 72

RESULT 15

Q86587 PRELIMINARY; PRT; 124 AA.
AC Q86587;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)
DE ORF 24 PROTEIN.
DE OS gallid herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10386;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93227567; PubMed=8385839;
RA Ross N., Binns M.M., Sanderson M., Schat K.A.;
RT "Alterations in DNA sequence and RNA transcription of the Bam HI-H
RT fragment accompany attenuation of Oncogenic Marek's disease
RT herpesvirus."
RL Virus Genes 7:33-51(1993).
DR EMBL: S58431; AAB26138.1; -;
SQ SEQUENCE 124 AA; 13963 MW; 389BAA029F177F30 CRC64;

Query Match 37.2%; Score 52.5; DB 12; Length 124;
Best Local Similarity 68.8%; Pred. No. 5.8;
Matches 11; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 4 WARGGAPRSGGRCRCR 19
I ||| I ||| I ||
Db 58 WKRGGPRGRRR-GR 72

Search completed: January 30, 2001, 16:55:06
Job time: 922 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:47:52 ; Search time 132.77 Seconds
(without alignments)
3.381 Million cell updates/sec

Title: US-09-518-931-4_COPY_142_166

Sequence: 141
1 GESWARGGAPRSGRRGRGVAGP 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep: *
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep: *
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep: *
4: /cgn2_6/ptodata/2/iaa/PCRU5_COMB.pep: *
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES				Description	
Result No.	Score	Query Match	Length DB ID		
1	55.5	39.4	181 2	US-08-726-306A-56	Sequence 56, Appli
2	53	37.6	461 1	US-08-194-338-4	Sequence 4, Appli
3	51	36.2	470 3	US-08-933-821-6	Sequence 6, Appli
4	51	36.2	470 3	US-08-960-507-6	Sequence 6, Appli
5	50.5	35.8	594 2	US-08-785-310A-6	Sequence 6, Appli
6	48	34.0	40 2	US-08-726-306A-48	Sequence 48, Appli
7	47.5	33.7	412 1	US-08-349-696-21	Sequence 21, Appli
8	47.5	33.7	412 1	US-08-233-009-21	Sequence 21, Appli
9	47.5	33.7	412 2	US-08-560-231-21	Sequence 21, Appli
10	47.5	33.7	412 3	US-08-080-704A-21	Sequence 21, Appli
11	46.5	33.0	1958 1	US-07-945-283-2	Sequence 2, Appli
12	46	32.6	223 1	US-07-667-276A-10	Sequence 10, Appli
13	46	32.6	581 3	US-09-135-021-80	Sequence 80, Appli
14	46	32.6	676 3	US-09-135-021-2	Sequence 2, Appli
15	45.5	32.3	714 2	US-08-990-114-3	Sequence 3, Appli
16	45	31.9	928 1	US-08-442-248-2	Sequence 2, Appli
17	45	31.9	928 1	US-08-440-815-2	Sequence 2, Appli
18	45	31.9	1005 2	US-08-469-537A-103	Sequence 103, App
19	45	31.9	1345 2	US-08-977-767-3	Sequence 3, Appli
20	44.5	31.6	53 3	US-08-897-924A-10	Sequence 10, Appli
21	44.5	31.6	56 3	US-08-897-924A-2	Sequence 2, Appli
22	44.5	31.6	58 3	US-08-897-924A-9	Sequence 9, Appli
23	44.5	31.6	67 3	US-08-897-924A-8	Sequence 8, Appli
24	44.5	31.6	210 1	US-08-464-590A-14	Sequence 14, Appli
25	44.5	31.6	210 2	US-08-207-412B-9	Sequence 9, Appli
26	44.5	31.6	210 3	US-09-093-585-14	Sequence 14, Appli
27	44.5	31.6	226 1	US-07-828-798C-7	Sequence 7, Appli
28	44.5	31.6	226 2	US-08-315-868A-7	Sequence 7, Appli

29 44.5 31.6 226 3 US-08-495-819B-7 Sequence 7, Appli
30 44.5 31.6 246 1 US-07-828-798C-6 Sequence 6, Appli
31 44.5 31.6 246 2 US-08-315-868A-6 Sequence 6, Appli
32 44.5 31.6 246 3 US-08-495-819B-6 Sequence 6, Appli
33 44.5 31.6 266 1 US-08-812-025-10 Sequence 10, Appli
34 44.5 31.6 266 3 US-07-791-931-10 Sequence 10, Appli
35 44 31.2 243 1 US-08-021-608D-6 Sequence 6, Appli
36 44 31.2 243 1 US-08-726-160-6 Sequence 6, Appli
37 44 31.2 243 4 PCT-US94-01782-6 Sequence 6, Appli
38 44 31.2 590 1 US-08-021-608D-8 Sequence 8, Appli
39 44 31.2 590 1 US-08-726-160-8 Sequence 8, Appli
40 44 31.2 590 4 PCT-US94-01782-8 Sequence 8, Appli
41 44 31.2 643 1 US-08-021-608D-10 Sequence 10, Appli
42 44 31.2 643 1 US-08-726-160-10 Sequence 10, Appli
43 44 31.2 643 4 PCT-US94-01782-10 Sequence 10, Appli
44 44 31.2 644 1 US-08-021-608D-2 Sequence 2, Appli
45 44 31.2 644 1 US-08-726-160-2 Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-726-306A-56
; Sequence 56, Application US/08726306A
; Patent No. 5958684
; GENERAL INFORMATION:
; APPLICANT: van Leeuwen, Frederik Willem
; APPLICANT: Burbach, Johannes Peter Henri
; TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS
; NUMBER OF SEQUENCES: 189
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 1 Financial Center
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726.306A
; FILING DATE: 02-Oct-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 95/20080.4
; FILING DATE: 02-Oct-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/009,832
; FILING DATE: 01-Jan-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Ph.D., Kathleen M.
; REGISTRATION NUMBER: 34,380
; REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 345-9100
; TELEFAX: (617) 345-9111
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 181 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-726-306A-56

Query Match 39.4%; Score 55.5; DB 2; Length 181;
Best Local Similarity 57.1%; Pred. No. 2.7;
Matches 12; Conservative 1; Mismatches 7; Indels 1; Gaps 1;

Query Match 36.2%; Score 51; DB 2; Length 470;
Best Local Similarity 81.8%; Pred. No. 23;
Matches 9; Conservative 0; Mismatches 2; Indels

```

RESULT      4
US-08-960-507-6
: Sequence 6, Application US/08960507
: Patent No. 6057435
: GENERAL INFORMATION:
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: TITLE OF INVENTION: Tie Ligands
: NUMBER OF SEQUENCES: 24
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genentech, Inc.
: STREET: 1 DNA Way

```

COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/960.507

FILING DATE: 11/13/01
 CLASSIFICATION: 536
 ATTORNEY/AGENT INFORMATION:
 NAME: Dreger, Ginger R.
 REGISTRATION NUMBER: 33,055
 REFERENCE/DOCKET NUMBER: p1130p1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650/225-3216
 TELEFAX: 650/952-9881
 INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
LENGTH: 470 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-960-507-6

Query Match 36.2%; Score 51; DB 3; Length 470;
Best Local Similarity 81.8%; Pred. No. 23;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GESWARGGAPR 11
I I I I I I I I I I
Db 16 GASWARAGAPR 26

RESULT 5
US-08-785-310A-6
; Sequence 6, Application US/08785310A
; Patent No. 5840532

GENERAL INFORMATION:
APPLICANT: McKnight, Steven L.
APPLICANT: Russell, David W.
TITLE OF INVENTION: Neuronal PAS Domain Protein
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/785,310A
FILING DATE: 21-JAN-1997

CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UTSD:1226
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 594 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-785-310A-6

Query Match 35.8%; Score 50.5; DB 2; Length 594;
Best Local Similarity 61.1%; Pred. No. 33;
Matches 11; Conservative 2; Mismatches 2; Indels 3; Gaps 1;

Qy 10 PRSGR---RCGRGVAG 24
I I I I I I I I I I I I
Db 6 PRSGRGVEKCGGRGAG 23

RESULT 6
US-08-726-306A-48
; Sequence 48, Application US/08726306A
; Patent No. 5958684
GENERAL INFORMATION:
APPLICANT: van Leeuwen, Frederik Willem

APPLICANT: Burbach, Johannes Peter Henri
APPLICANT: Grosveld, Franklin G.
TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS
NUMBER OF SEQUENCES: 189
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff, Ltd.
STREET: 1 Financial Center
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/726,306A
FILING DATE: 02-Oct-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 95/20080.4
FILING DATE: 02-Oct-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/009,832
FILING DATE: 01-Jan-1996
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Ph.D., Kathleen M.
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 345-9100
TELEFAX: (617) 345-9111
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-726-306A-48

Query Match 34.0%; Score 48; DB 2; Length 40;
Best Local Similarity 47.6%; Pred. No. 5.9;
Matches 10; Conservative 2; Mismatches 7; Indels 2; Gaps 1;

Qy 4 WARG--GAPRSGRRRCGRGV 22
I I I I I I I I I I I I
Db 2 WGRGKEGMQEGGRRRGSKI 22

RESULT 7
US-08-349-696-21
; Sequence 21, Application US/08349696
; Patent No. 5599671
GENERAL INFORMATION:
APPLICANT: Jacobson, Marlene A
APPLICANT: Johnson, Robert G
APPLICANT: Luneau, Christopher J
APPLICANT: Salvatore, Christopher A
TITLE OF INVENTION: Human Adenosine Receptors
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000
CITY: Rahway
STATE: NJ
COUNTRY: United States
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh IIfx
OPERATING SYSTEM: Macintosh

SOFTWARE: Microsoft Word 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/349,696
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: us/08/005945
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Meredith, Roy D.
REGISTRATION NUMBER: 30,777
REFERENCE/DOCKET NUMBER: 186991A
TELEPHONE: (908)594-4678
TELEFAX: (908)594-4720
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 412 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-349-696-21

Query Match 33.7%; Score 47.5; DB 1; Length 412;
Best Local Similarity 43.5%; Pred. No. 56;
Matches 10; Conservative 2; Mismatches 8; Indels 3; Gaps 1;

QY 4 WARGGAPRSG---GRRCGRGQVA 23
| | | | | : | | | | |
Db 143 WNNCGQPKGKNGHNSQCGGQVA 165

RESULT 8

US-08-233-009-21
Sequence 21, Application US/082333009
Patent No. 5646156
GENERAL INFORMATION:
APPLICANT: Jacobson, Marlene A
APPLICANT: Johnson, Robert G
APPLICANT: Salvatore, Christopher A
TITLE OF INVENTION: INHIBITION OF EOSINOPHIL
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000
CITY: Rahway
STATE: New Jersey
COUNTRY: United States
ZIP: 07065
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/233,009
FILING DATE: 25-APR-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Bencen, Gerard H
REGISTRATION NUMBER: 35,746
REFERENCE/DOCKET NUMBER: 19219
TELEPHONE: (908) 594-3901
TELEFAX: (908)594-4720
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 412 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-233-009-21

Query Match 33.7%; Score 47.5; DB 1; Length 412;
Best Local Similarity 43.5%; Pred. No. 56;
Matches 10; Conservative 2; Mismatches 8; Indels 3; Gaps 1;

QY 4 WARGGAPRSG---GRRCGRGQVA 23
| | | | | : | | | | |
Db 143 WNNCGQPKGKNGHNSQCGGQVA 165

RESULT 9

US-08-560-231-21
Sequence 21, Application US/08560231
Patent No. 5817760
GENERAL INFORMATION:
APPLICANT: Jacobson, Marlene A
APPLICANT: Johnson, Robert G
APPLICANT: Luneau, Christopher J
APPLICANT: Salvatore, Christopher A
TITLE OF INVENTION: Human Adenosine Receptors
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: P.O. Box 2000
CITY: Rahway
STATE: NJ
COUNTRY: United States
ZIP: 07065
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh IIcx
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,231
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Meredith, Roy D.
REGISTRATION NUMBER: 30,777
REFERENCE/DOCKET NUMBER: 186991A
TELEPHONE: (908)594-4678
TELEFAX: (908)594-4720
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 412 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-560-231-21

Query Match 33.7%; Score 47.5; DB 2; Length 412;
Best Local Similarity 43.5%; Pred. No. 56;
Matches 10; Conservative 2; Mismatches 8; Indels 3; Gaps 1;

QY 4 WARGGAPRSG---GRRCGRGQVA 23
| | | | | : | | | | |
Db 143 WNNCGQPKGKNGHNSQCGGQVA 165

RESULT 10

US-09-080-704A-21
Sequence 21, Application US/09080704A
Patent No. 6166181

```
;
;
; GENERAL INFORMATION:
; APPLICANT: Jacobson, Marlene A
; APPLICANT: Johnson, Robert G
; APPLICANT: Luneau, Christopher J
; APPLICANT: Salvatore, Christopher A
; TITLE OF INVENTION: Human Adenosine Receptors
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: P.O. Box 2000
; CITY: Rahway
; STATE: NJ
; COUNTRY: United States
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/080,704A
; FILING DATE: 18 May 1998
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Parr, Richard S.
; REGISTRATION NUMBER: 32,586
; REFERENCE/DOCKET NUMBER: 18699DB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (732)594-4958
; TELEFAX: (732)594-4720
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 412 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-080-704A-21

Query Match 33.7%; Score 47.5; DB 3; Length 412;
Best Local Similarity 43.5%; Pred. No. 56;
Matches 10; Conservative 2; Mismatches 8; Indels 3; Gaps 1;

Qy 4 WARGAPRSG---GRCRCGOVA 23
Db 143 WNCGPQPKGNHSGCGEGQVA 165

RESULT 11
; Sequence 2, Application US/07945283
; Patent No. 5352596
; GENERAL INFORMATION:
; APPLICANT: Cheung, Andrew K.
; APPLICANT: Wesley, Ronald D.
; TITLE OF INVENTION: Pseudorabies Virus Deletion Mutants
; TITLE OF INVENTION: Involving The EP0 and LLT Genes
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtiss P. Ribando
; STREET: 1815 No. 5352596th University Street
; CITY: Peoria
; STATE: IL
; COUNTRY: USA
; ZIP: 61604
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/945,283
```

```
;
;
; FILING DATE: 19920911
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Ribando, Curtiss P
; REGISTRATION NUMBER: 27976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 309-685-4011 ext.513
; TELEFAX: 309-685-4128
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1958 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-945-283-2

Query Match 33.0%; Score 46.5; DB 1; Length 1958;
Best Local Similarity 62.5%; Pred. No. 3.1e+02;
Matches 10; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

Qy 7 GGAPRSGGRRRCRGQV 22
Db 1683 GGGP-GGGGRAGRGV 1697

RESULT 12
; US-07-667-276A-10
; Sequence 10, Application US/07667276A
; Patent No. 5470971
; GENERAL INFORMATION:
; APPLICANT: Kondo, Keiji
; APPLICANT: Inouye, Masayori
; TITLE OF INVENTION: STRESS-INDUCED PROTEINS, GENES CODING
; TITLE OF INVENTION: THEREFOR, TRANSFORMED CELLS OF ORGANISMS, METHODS AND
; TITLE OF INVENTION: APPLICATIONS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Weiser & Associates
; STREET: 230 S. Fifteenth Street, Suite 500
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/667,276A
; FILING DATE: 11-MAR-1991
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiser, Gerard J.
; REGISTRATION NUMBER: 19,763
; REFERENCE/DOCKET NUMBER: 377,5351P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-875-8383
; TELEFAX: 215-875-8394
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-07-667-276A-10

Query Match 32.6%; Score 46; DB 1; Length 223;
Best Local Similarity 45.8%; Pred. No. 50;
Matches 11; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
```

QY 1 GESWARGGAPRRGGRGGRGVAG 24
| | | | | | | | | |
Db 168 GRGGGGRGGGGRGGGGRGGFGG 191

RESULT 13

US-09-135-021-80
; Sequence 80, Application US/09135021A
; Patent No. 6150104
; GENERAL INFORMATION:
; APPLICANT: Splawski, Igor
; APPLICANT: Keating, Mark T.
; TITLE OF INVENTION: A HOMOZYGOUS MUTATION IN KVLQT1 WHICH CAUSES JERVELL
; TITLE OF INVENTION: AND LANGE-NIELSEN SYNDROME
; FILE REFERENCE: 2323-128
; CURRENT APPLICATION NUMBER: US/09/135,021A
; CURRENT FILING DATE: 1998-08-17
; EARLIER APPLICATION NUMBER: 08/874,655
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/094,477
; EARLIER FILING DATE: 1998-07-29
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 80
; LENGTH: 581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-135-021-80

Query Match 32.6%; Score 46; DB 3; Length 581;
Best Local Similarity 37.9%; Pred. No. 1.2e+02;
Matches 11; Conservative 2; Mismatches 12; Indels 4; Gaps 1;

QY 1 GESWARGGAPRRSGG----RRCGRGVAGP 25
| : | | | | | : | | | | |
Db 527 GSTPGSGGPPREGGAHITPCGSGGSVDP 555

RESULT 14

US-09-135-021-2
; Sequence 2, Application US/09135021A
; Patent No. 6150104
; GENERAL INFORMATION:
; APPLICANT: Splawski, Igor
; APPLICANT: Keating, Mark T.
; TITLE OF INVENTION: A HOMOZYGOUS MUTATION IN KVLQT1 WHICH CAUSES JERVELL
; TITLE OF INVENTION: AND LANGE-NIELSEN SYNDROME
; FILE REFERENCE: 2323-128
; CURRENT APPLICATION NUMBER: US/09/135,021A
; CURRENT FILING DATE: 1998-08-17
; EARLIER APPLICATION NUMBER: 08/874,655
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/094,477
; EARLIER FILING DATE: 1998-07-29
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 676
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-135-021-2

Query Match 32.6%; Score 46; DB 3; Length 676;
Best Local Similarity 37.9%; Pred. No. 1.4e+02;
Matches 11; Conservative 2; Mismatches 12; Indels 4; Gaps 1;

QY 1 GESWARGGAPRRSGG----RRCGRGVAGP 25
| : | | | | | : | | | | |
Db 622 GSTPGSGGPPREGGAHITPCGSGGSVDP 650

RESULT 15
US-08-990-114-3
; Sequence 3, Application US/08990114
; Patent No. 5932475
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN NUCLEOLIN-LIKE PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/990.114
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0451 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 714 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 128842
US-08-990-114-3

Query Match 32.3%; Score 45.5; DB 2; Length 714;
Best Local Similarity 48.0%; Pred. No. 1.6e+02;
Matches 12; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

QY 1 GESWARGG-APRRGGRGGRGVAG 24
| | | | | | | | | |
Db 655 GRGGGGRGGGGRGGGGRGGFGG 679

Search completed: January 30, 2001, 16:47:53
Job time: 629 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:47:59 ; Search time 29.47 Seconds
(without alignments)
348.087 Million cell updates/sec

Title: US-09-518-931-2
Perfect score: 300
Sequence: 1 MRALGPGSLLLCLVLALPA.....RVARMPGLSRRERFLPVH 300

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 268485 seqs, 34193795 residues

Word size : 6
Total number of hits satisfying chosen parameters: 1259

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : A_Geneseq_36.*
1: /SIDS1/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	300	100.0	300	19	W66102 Amino acid sequenc
2	300	100.0	300	19	W63622 Human tumour necro
3	300	100.0	300	20	Y03099 Human lung TNF-rec
4	300	100.0	300	20	Y42182 Human FLINT #1 pro
5	300	100.0	300	20	Y17479 Mammalian tumour n
6	300	100.0	300	20	Y06817 Human DcR3 polypep
7	300	100.0	300	20	W97749 Human tumour necro
8	300	100.0	300	20	W95082 Orphan receptor (H
9	271	90.3	271	20	Y42184 Human mFLINT #1 pr
10	240	80.0	245	20	Y28449 A human tumour nec
11	215	71.7	215	20	W93585 Human hAPO6 protei
12	208	69.3	300	21	Y77458 Human TNF receptor

Human TNFR superfa	153	51.0	20	Y22222	Human TNFR superfa		
Human tumour necro	142	47.3	19	W63623	Human tumour necro		
Human mFLINT #2 pr	130	43.3	20	Y42185	Human mFLINT #2 pr		
Human FLINT #2 pro	130	43.3	20	Y42183	Human FLINT #2 pro		
Human TANGO 129 (T	17	9	3.0	408	20	Y41111	Human TANGO 129 (T
Human TANGO 129 (T	18	9	3.0	430	20	Y41110	Human TANGO 129 (T
Human tumour necro	19	9	3.0	430	21	Y70785	Human tumour necro
Rat HFGAN72 recept	20	8	2.7	20	Y24805	Rat HFGAN72 recept	
Mouse HFGAN72 rece	21	8	2.7	20	Y24808	Mouse HFGAN72 rece	
Human HFGAN72 rece	22	8	2.7	20	Y24801	Human HFGAN72 rece	
TNFR/NGFR cysteine	23	8	2.7	40	20	Y41113	TNFR/NGFR cysteine
Mouse HFGAN72 rece	24	8	2.7	123	19	W61383	Mouse HFGAN72 rece
Mouse HFGAN72 rece	25	8	2.7	123	20	Y24806	Mouse HFGAN72 rece
Rat HFGAN72 recept	26	8	2.7	130	19	W61382	Rat HFGAN72 recept
Rat hypocrerin 35	27	8	2.7	130	19	W50157	Rat hypocrerin 35
Mouse hypocrerin 3	28	8	2.7	130	19	W50158	Mouse hypocrerin 3
Rat HFGAN72 recept	29	8	2.7	130	20	Y24802	Rat HFGAN72 recept
Human HFGAN72 rece	30	8	2.7	131	19	W61381	Human HFGAN72 rece
Human HFGAN72 rece	31	8	2.7	131	20	Y24799	Human HFGAN72 rece
Tylactone synthase	32	8	2.7	472	18	W22601	Tylactone synthase
Secretin precursor	33	7	2.3	10	3	P20401	Secretin precursor
Intermediate in se	34	7	2.3	10	4	P30018	Intermediate in se
Estriadiol activat	35	7	2.3	16	21	Y65459	Estriadiol activat
Secretin precursor	36	7	2.3	17	3	P20400	Secretin precursor
Intermediate of se	37	7	2.3	17	4	P30019	Intermediate of se
Protected heptacos	38	7	2.3	27	3	P20383	Protected heptacos
Secretin precursor	39	7	2.3	27	3	P20398	Secretin precursor
Sequence of 27-Dea	40	7	2.3	27	4	P30551	Sequence of 27-Dea
Synthetic secretin	41	7	2.3	27	4	P30021	Synthetic secretin
Intermediate in se	42	7	2.3	27	4	P30049	Intermediate in se
27-Desamidosecreti	43	7	2.3	27	4	P30014	27-Desamidosecreti
Pig Secretin. Sus	44	7	2.3	27	4	P30038	Pig Secretin. Sus
Secretin protein s	45	7	2.3	27	7	P60847	Secretin protein s
Human glucagon deg	46	7	2.3	17	17	R93024	Human glucagon deg
Secretin-derived t	47	7	2.3	27	19	W71676	Secretin-derived t
Porcine secretin p	48	7	2.3	27	19	W37793	Porcine secretin p
Rabbit secretin pe	49	7	2.3	27	19	W37794	Rabbit secretin pe
Porcine secretin p	50	7	2.3	27	19	W37795	Porcine secretin p
Porcine secretin p	51	7	2.3	27	19	W37796	Porcine secretin p
Neutrophil-activat	52	7	2.3	27	20	Y02336	Neutrophil-activat
Recombinantly prod	53	7	2.3	28	4	P30063	Recombinantly prod
27-desamidosecreti	54	7	2.3	28	4	P30062	27-desamidosecreti
Human secretin pre	55	7	2.3	28	10	P91869	Human secretin pre
Human secretin. H	56	7	2.3	31	10	P90130	Human secretin. H
Sequence encoded b	57	7	2.3	33	8	P70421	Sequence encoded b
HSV-1 ICP4 serine	58	7	2.3	69	18	W22730	HSV-1 ICP4 serine
Amino acid sequenc	59	7	2.3	94	20	Y35734	Amino acid sequenc
Amino acid sequenc	60	7	2.3	108	20	Y10879	Amino acid sequenc
652 strain leukoto	61	7	2.3	150	19	W41364	652 strain leukoto
A.actinomycetemcom	62	7	2.3	150	20	Y02643	A.actinomycetemcom
Breast cancer asso	63	7	2.3	155	20	Y07014	Breast cancer asso
Glycosylphosphatid	64	7	2.3	160	16	R79056	Glycosylphosphatid
Soluble herpesviru	65	7	2.3	186	21	Y79206	Soluble herpesviru
Soluble herpesviru	66	7	2.3	193	21	Y79204	Soluble herpesviru
Soluble herpesviru	67	7	2.3	197	21	Y79205	Soluble herpesviru
Amino acid sequenc	68	7	2.3	207	20	Y29151	Amino acid sequenc
Human Apo B protel	69	7	2.3	208	20	Y31236	Human Apo B protel
Glycosylphosphatid	70	7	2.3	233	16	R79057	Glycosylphosphatid
H. pylori GHPO 127	71	7	2.3	277	19	W98512	H. pylori GHPO 127
Membrane-bound her	72	7	2.3	277	21	Y79207	Membrane-bound her
Human tumour necro	73	7	2.3	283	17	W05809	Human tumour necro
Human herpes simpl	74	7	2.3	283	18	W12659	Human herpes simpl
Herpesvirus entry	75	7	2.3	283	19	W62338	Herpesvirus entry
Human TNF receptor	76	7	2.3	283	19	W60045	Human TNF receptor
Human tumour-assoc	77	7	2.3	283	20	Y06488	Human tumour-assoc
Herpes virus entry	78	7	2.3	283	20	Y05797	Herpes virus entry
Tumour necrosis fa	79	7	2.3	283	20	W95031	Tumour necrosis fa
Human tumour necro	80	7	2.3	283	20	W87591	Human tumour necro
Sequence of the N-	81	7	2.3	291	8	P70659	Sequence of the N-
Homo sapiens sprou	82	7	2.3	315	19	W48792	Homo sapiens sprou
Amino acid sequenc	83	7	2.3	352	20	Y35164	Amino acid sequenc
Phaffia derived ca	84	7	2.3	376	18	W22498	Phaffia derived ca
Cps2G protein whic	85	7	2.3	385	21	Y68959	Cps2G protein whic

86	7	2.3	409	15	R55790	Cell death hid pro	159	6	2.0	26	20	Y05650
87	7	2.3	410	15	R55791	Cell death hid pro	160	6	2.0	31	18	W17085
88	7	2.3	411	21	Y94423	Drosophila melanog	161	6	2.0	35	19	Y21447
89	7	2.3	410	21	Y94424	Drosophila melanog	162	6	2.0	36	7	P60278
90	7	2.3	414	20	Y07091	Colon cancer assoc	163	6	2.0	37	20	W88882
91	7	2.3	430	19	W64720	Human TIABP2 (TIAK	164	6	2.0	38	15	R58398
92	7	2.3	432	20	Z36947	Amino acid sequenc	165	6	2.0	42	20	Y27638
93	7	2.3	435	8	P70661	Sequence of fusion	166	6	2.0	42	20	Y11899
94	7	2.3	447	8	P70660	Sequence of the N-	167	6	2.0	43	20	Y25807
95	7	2.3	467	21	Y54602	AV37 antigen prote	168	6	2.0	43	20	Y31323
96	7	2.3	470	20	Y24314	Human repressor Kr	169	6	2.0	43	20	Y31324
97	7	2.3	479	20	Y24315	Human repressor Kr	170	6	2.0	43	20	Y31325
98	7	2.3	493	20	Y07096	Colon cancer assoc	171	6	2.0	43	20	Y31309
99	7	2.3	515	18	W31547	Bloom's syndrome B	172	6	2.0	43	20	Y31310
100	7	2.3	543	18	W07702	Mouse ETS2 repress	173	6	2.0	43	20	Y31311
101	7	2.3	548	18	W07700	Human ETS2 repress	174	6	2.0	43	20	Y31312
102	7	2.3	568	20	Y37828	Chlamydia trachoma	175	6	2.0	43	20	Y31313
103	7	2.3	590	19	W62832	Gossypium hirsutum	176	6	2.0	43	20	Y31314
104	7	2.3	615	19	W29672	Homo sapiens clone	177	6	2.0	43	20	Y31315
105	7	2.3	655	21	Y59207	B. halodurans clon	178	6	2.0	43	20	Y31316
106	7	2.3	671	21	Y99426	Human PRO1604 (UNQ	179	6	2.0	43	20	Y31317
107	7	2.3	737	16	R76640	Deltex protein, UNQ	180	6	2.0	43	20	Y31318
108	7	2.3	737	16	R76639	Deltex protein pro	181	6	2.0	43	20	Y31319
109	7	2.3	737	18	W18317	Drosophila Deltex	182	6	2.0	43	20	Y31320
110	7	2.3	739	18	W31549	Bloom's syndrome B	183	6	2.0	43	20	Y31321
111	7	2.3	754	20	Y07027	Breast cancer asso	184	6	2.0	43	20	Y31322
112	7	2.3	809	20	Y09371	Human apolipoprote	185	6	2.0	43	20	Y12205
113	7	2.3	812	20	Y52775	Treponema pallidum	186	6	2.0	44	18	Y11159
114	7	2.3	837	20	Y52774	Treponema pallidum	187	6	2.0	44	20	W97674
115	7	2.3	845	16	R70065	Hepatitis B virus	188	6	2.0	45	21	Y59061
116	7	2.3	849	17	W06725	FLO1 protein, invo	189	6	2.0	50	20	Y27662
117	7	2.3	875	19	W37098	Murine endothelial	190	6	2.0	52	20	W67819
118	7	2.3	915	20	Y13350	Amino acid sequenc	191	6	2.0	54	16	R71472
119	7	2.3	950	20	W83319	Mouse SMAD interac	192	6	2.0	55	20	Y11426
120	7	2.3	971	18	W11631	Human cellular apo	193	6	2.0	57	20	Y13123
121	7	2.3	1035	20	Y49128	Chimeric hmgIuR2/h	194	6	2.0	58	20	W67821
122	7	2.3	1093	16	R66460	AF-17 protein, Ho	195	6	2.0	59	18	W28271
123	7	2.3	1136	20	Y23330	Human tumour suppr	196	6	2.0	62	20	Y12367
124	7	2.3	1136	20	Y22655	Human tumour suppr	197	6	2.0	67	21	Y52254
125	7	2.3	1167	20	Y42709	Arabidopsis thalia	198	6	2.0	69	20	Y11894
126	7	2.3	1173	20	Y42710	Arabidopsis thalia	199	6	2.0	70	21	Y75953
127	7	2.3	1173	20	Y42714	Arabidopsis thalia	200	6	2.0	72	20	Y48285
128	7	2.3	1297	19	W40200	Infected cell prot	201	6	2.0	87	13	R22386
129	7	2.3	1298	19	W80810	The amino acid seq	202	6	2.0	87	20	W88402
130	7	2.3	1394	20	Y49129	pmGluR2/Car*Galpha	203	6	2.0	92	18	W41642
131	7	2.3	1397	20	Y49134	pmGluR2/Car*Galpha	204	6	2.0	92	18	W41643
132	7	2.3	1417	18	W31548	Bloom's syndrome B	205	6	2.0	92	21	Y69955
133	7	2.3	1417	18	W31550	Bloom's syndrome B	206	6	2.0	94	20	Y02936
134	7	2.3	1417	18	W31551	Bloom's syndrome B	207	6	2.0	96	20	Y12206
135	7	2.3	1418	18	W15264	Bloom syndrome act	208	6	2.0	98	12	R14681
136	7	2.3	1611	18	W22604	Ty lactone synthase	209	6	2.0	101	16	R77251
137	7	2.3	1671	11	R06341	Signal peptide and	210	6	2.0	103	16	R80091
138	7	2.3	1798	21	Y51611	Human HSGT1 protei	211	6	2.0	103	20	W95489
139	7	2.3	2721	8	P70647	Sequence of N-term	212	6	2.0	104	20	Y12207
140	7	2.3	3910	14	R38470	ALL-1 protein, Ho	213	6	2.0	106	19	Y20986
141	7	2.3	3910	16	R66462	ALL-1 (acute lymph	214	6	2.0	107	20	Y11832
142	7	2.3	3969	15	R52971	Product of the cDN	215	6	2.0	112	20	Y09279
143	6	2.0	7	18	W45347	Peptide #3 bearing	216	6	2.0	117	20	Y73877
144	6	2.0	7	21	Y60941	Cadherin-5 cell ad	217	6	2.0	118	13	R27755
145	6	2.0	8	21	Y60942	Cadherin-5 cell ad	218	6	2.0	118	18	W20129
146	6	2.0	9	18	W10275	Antiphospholipid a	219	6	2.0	120	20	Y76543
147	6	2.0	9	19	W67101	Antibody ACA-6501	220	6	2.0	122	15	R63207
148	6	2.0	9	19	W46066	Antiphospholipid a	221	6	2.0	122	16	R85472
149	6	2.0	9	21	Y60943	Cadherin-5 cell ad	222	6	2.0	122	20	Y25614
150	6	2.0	10	19	W82809	Cytomodulating lip	223	6	2.0	124	18	W22392
151	6	2.0	10	21	Y60946	Cadherin-5 cell ad	224	6	2.0	124	21	Y49943
152	6	2.0	11	21	Y60949	Cadherin-5 cell ad	225	6	2.0	125	21	R94670
153	6	2.0	13	18	W45349	Peptide #6 bearing	226	6	2.0	126	17	R91220
154	6	2.0	13	18	W30936	Birch pollen aller	227	6	2.0	126	18	W31381
155	6	2.0	16	18	W31042	Mugwort pollen all	228	6	2.0	126	20	Y41216
156	6	2.0	20	14	R38543	Foot and Mouth Dis	229	6	2.0	126	20	W97223
157	6	2.0	20	14	R14793	Foot and Mouth Dis	230	6	2.0	126	20	W95183
158	6	2.0	22	18	W45350	Peptide #7 bearing	231	6	2.0	127	17	Y98891

232	6	2.0	127	21	Y95025	Human clone vc51.1	305	6	2.0	174	13	R25427	Recombinant G-CSF
233	6	2.0	128	21	Y73489	Human secreted pro	306	6	2.0	174	14	R34707	Human CSF analogue
234	6	2.0	130	18	W33893	Flea saliva protei	307	6	2.0	174	15	R56568	G-CSF analogue, De
235	6	2.0	130	19	W82367	Flea saliva protei	308	6	2.0	174	16	R75332	Human granulocyte-
236	6	2.0	132	19	W75070	Human secreted pro	309	6	2.0	174	16	R68626	Human G-CSF. Homo
237	6	2.0	133	20	Y27816	Human secreted pro	310	6	2.0	174	17	W02204	Human G-CSF for st
238	6	2.0	135	20	Y29154	Amino acid sequenc	311	6	2.0	174	18	W17090	Generic backbone o
239	6	2.0	136	17	R94879	Actinomadura flexu	312	6	2.0	174	18	W15025	G-CSF receptor ago
240	6	2.0	137	20	Y05945	Thermophilus therm	313	6	2.0	174	18	W15034	G-CSF receptor ago
241	6	2.0	139	20	Y41245	Human Y218 protein	314	6	2.0	174	18	W15035	G-CSF receptor ago
242	6	2.0	141	21	Y73438	Human secreted pro	315	6	2.0	174	18	W15036	G-CSF receptor ago
243	6	2.0	142	16	R75758	BAV3 ORF3 product.	316	6	2.0	174	18	W15037	G-CSF receptor ago
244	6	2.0	142	18	Y11260	Streptococcus pneu	317	6	2.0	174	18	W15038	G-CSF receptor ago
245	6	2.0	143	16	R85476	PTLV-L Rof1 protei	318	6	2.0	174	18	W15039	G-CSF receptor ago
246	6	2.0	143	16	R76951	Fusion protein of	319	6	2.0	174	18	W15040	G-CSF receptor ago
247	6	2.0	144	21	Y70582	Salmonella Pathoge	320	6	2.0	174	18	W15041	G-CSF receptor ago
248	6	2.0	146	20	Y59845	Human normal uteru	321	6	2.0	174	18	W15042	G-CSF receptor ago
249	6	2.0	148	18	W12693	Human cystatin E.	322	6	2.0	174	18	W15052	G-CSF receptor ago
250	6	2.0	149	18	W15791	Human cysteine pro	323	6	2.0	174	18	W15053	G-CSF receptor ago
251	6	2.0	149	19	W55826	Cysteine protease	324	6	2.0	174	18	W15054	G-CSF receptor ago
252	6	2.0	149	20	Y74142	Human prostate tum	325	6	2.0	174	18	W15055	G-CSF receptor ago
253	6	2.0	149	21	Y67351	Human cystatin E (326	6	2.0	174	18	W15057	G-CSF receptor ago
254	6	2.0	149	21	Y49912	Human cystatin E.	327	6	2.0	174	18	W15058	G-CSF receptor ago
255	6	2.0	150	20	Y73849	Human prostate tum	328	6	2.0	174	18	W15059	G-CSF receptor ago
256	6	2.0	150	20	Y23765	A human phosphodie	329	6	2.0	174	18	W15060	G-CSF receptor ago
257	6	2.0	151	19	W38678	S. pneumoniae SPOO	330	6	2.0	174	18	W01606	Granulocyte-colony
258	6	2.0	154	12	R14358	Prod. of acetic ac	331	6	2.0	174	18	W01607	Granulocyte-colony
259	6	2.0	154	20	Y37855	Amino acid sequenc	332	6	2.0	174	18	W01608	Granulocyte-colony
260	6	2.0	156	14	R40840	Glioblastoma deriv	333	6	2.0	174	19	W77783	Modified human gra
261	6	2.0	158	19	R86114	S. pneumoniae deri	334	6	2.0	174	21	Y82680	Human granulocyte
262	6	2.0	159	15	R56561	G-CSF analogue, K1	335	6	2.0	174	21	Y78937	Granulocyte colony
263	6	2.0	162	17	R98892	Recombinant plant	336	6	2.0	175	8	P70732	Sequence encoded b
264	6	2.0	166	19	W98727	H. pylori GHPO 892	337	6	2.0	175	8	P71030	Sequence of human
265	6	2.0	167	20	Y37335	Amino acid sequenc	338	6	2.0	175	10	P91070	Human granulocyte
266	6	2.0	169	16	R78612	Plasmid fragment p	339	6	2.0	175	10	P90170	Human granulocyte
267	6	2.0	169	18	W15044	G-CSF receptor ago	340	6	2.0	175	10	P90107	Gly 17 human granu
268	6	2.0	171	18	W15043	G-CSF receptor ago	341	6	2.0	175	11	R04898	Ala 17 human granu
269	6	2.0	171	18	W15045	G-CSF receptor ago	342	6	2.0	175	11	R06421	Ser[17,27]G-CSF
270	6	2.0	171	18	W15046	G-CSF receptor ago	343	6	2.0	175	12	R14224	G-CSF mutant L51X.
271	6	2.0	171	18	W15047	G-CSF receptor ago	344	6	2.0	175	12	R14730	G-CSF mutant L51X.
272	6	2.0	171	18	W15048	G-CSF receptor ago	345	6	2.0	175	12	R14731	G-CSF mutant G52X.
273	6	2.0	171	18	W15049	G-CSF receptor ago	346	6	2.0	175	12	R14732	G-CSF mutant H53X.
274	6	2.0	171	18	W15050	G-CSF receptor ago	347	6	2.0	175	12	R14733	G-CSF mutant S54X.
275	6	2.0	171	18	W15051	G-CSF receptor ago	348	6	2.0	175	12	R14739	G-CSF mutant L55X.
276	6	2.0	174	8	P70023	Modified granulocy	349	6	2.0	175	12	R14734	G-CSF mutant G56X.
277	6	2.0	174	9	P82483	Human G-CSF deriv.	350	6	2.0	175	12	R14735	G-CSF mutant H40X.
278	6	2.0	174	9	P80965	G-CSF deduced from	351	6	2.0	175	12	R14736	G-CSF mutant H157X.
279	6	2.0	174	10	P90470	G-colony stimulat	352	6	2.0	175	12	R14737	Cancer metastasis
280	6	2.0	174	10	P90501	ND28 deriv. of G-C	353	6	2.0	175	12	R14738	G-CSF analogue, D2
281	6	2.0	174	11	R04887	Derivative of huma	354	6	2.0	175	15	R56576	G-CSF analogue, M1
282	6	2.0	174	11	R05113	hg-CSF[ND28] as en	355	6	2.0	175	15	R56577	G-CSF analogue, M1
283	6	2.0	174	11	R05115	hg-CSF[ND28N145] a	356	6	2.0	175	15	R56578	G-CSF analogue, M1
284	6	2.0	174	11	R05114	hg-CSF[ND28N6] as	357	6	2.0	175	15	R56575	G-CSF analogue, C1
285	6	2.0	174	11	R05984	Modified human gra	358	6	2.0	175	15	R56576	G-CSF analogue, Q1
286	6	2.0	174	11	R07457	Protein having gra	359	6	2.0	175	15	R56577	G-CSF analogue, Q1
287	6	2.0	174	11	R07664	Cysteine-added var	360	6	2.0	175	15	R56582	G-CSF analogue, E2
288	6	2.0	174	12	R13679	Granulocyte colony	361	6	2.0	175	15	R56584	G-CSF analogue, M1
289	6	2.0	174	12	R15203	[Ser 17,27]huG-CSF	362	6	2.0	175	15	R56579	G-CSF analogue, M1
290	6	2.0	174	12	R15204	[Arg11,Ser17,27,60	363	6	2.0	175	15	R56580	G-CSF analogue, Q1
291	6	2.0	174	12	R15205	[Glu15,Ser17,27,Al	364	6	2.0	175	15	R56582	G-CSF analogue, Q1
292	6	2.0	174	12	R15206	[Arg11,Glu15,Ser17	365	6	2.0	175	15	R56583	G-CSF analogue, E2
293	6	2.0	174	12	R15207	[Arg11,165,Glu15,S	366	6	2.0	175	15	R56584	G-CSF analogue, M1
294	6	2.0	174	12	R15208	[Arg11,23,Ser17,27	367	6	2.0	175	15	R56585	G-CSF analogue, S1
295	6	2.0	174	12	R15209	[Arg11,40,Ser17,27	368	6	2.0	175	15	R56586	G-CSF analogue, K1
296	6	2.0	174	12	R15210	[Glu15,I11,Ser17,2	369	6	2.0	175	15	R56587	G-CSF analogue, Q2
297	6	2.0	174	12	R15211	[Ala1,Thr3,Tyr4,Ar	370	6	2.0	175	15	R56588	G-CSF analogue, Q2
298	6	2.0	174	12	R15212	[Glu15,Ser17,27,Al	371	6	2.0	175	15	R56589	G-CSF analogue, H4
299	6	2.0	174	12	R15213	[Ser17,27,60,65]hu	372	6	2.0	175	15	R56590	G-CSF analogue, H5
300	6	2.0	174	12	R15214	[Arg11,Ser17,27,65	373	6	2.0	175	15	R56591	G-CSF analogue, D1
301	6	2.0	174	12	R15215	[Ser17,27,65]huG-C	374	6	2.0	175	15	R56592	G-CSF analogue, D1
302	6	2.0	174	13	R26909	Human G-CSF. synt	375	6	2.0	175	15	R56593	G-CSF analogue, T1
303	6	2.0	174	13	R26910	Human (Ser17,27) G	376	6	2.0	175	15	R56594	G-CSF analogue, D2
304	6	2.0	174	13	R23595	Recombinant hemato	377	6	2.0	175	15	R56595	G-CSF analogue, E1

378	175	15	R56596	G-CSF analogue, F1	451	207	8	P70161	Human G-CSF encode
379	175	15	R56559	G-CSF analogue, K1	452	207	8	P71383	Sequence of human
380	175	15	R56560	G-CSF analogue, K2	453	207	8	P70163	Human G-CSF deduce
381	175	15	R56562	G-CSF analogue, O6	454	207	10	P90051	Human granulocyte
382	175	15	R56563	G-CSF analogue, C3	455	207	11	R05085	Mammalian amino ac
383	175	15	R56564	G-CSF analogue, Q2	456	211	20	X31151	Human XPC protein
384	175	15	R56565	G-CSF analogue, Q1	457	211	20	X31150	Murine XPC protein
385	175	15	R56566	G-CSF analogue, R1	458	212	12	R12521	B cell differentia
386	175	15	R56567	G-CSF analogue, R1	459	214	20	X37296	Amino acid sequenc
387	175	15	R56568	G-CSF analogue, K4	460	214	20	X37296	Human secreted pro
388	175	15	R56570	G-CSF analogue, H4	461	215	17	R98890	Recombinant transc
389	175	15	R56571	G-CSF analogue, E4	462	218	17	R98893	Coat protein of th
390	175	15	R56572	G-CSF analogue, R2	463	218	17	R98895	Coat protein of th
391	175	15	R56573	G-CSF analogue, K2	464	218	19	W59198	Seq ID 92 from US
392	175	15	R56574	G-CSF analogue, E2	465	218	19	W40087	Seq ID 92 from US
393	175	15	R56553	G-CSF analogue, K1	466	218	20	X37296	Amino acid sequenc
394	175	15	R56554	G-CSF analogue, K2	467	219	17	R77660	HPV E6-binding pro
395	175	15	R56555	G-CSF analogue, K3	468	219	19	W59197	Seq ID 91 from US
396	175	15	R56556	G-CSF analogue, K4	469	219	19	W40086	Seq ID 91 from US
397	175	15	R56557	G-CSF analogue, K1	470	221	20	X45219	Human CASB47 prote
398	175	15	R56558	G-CSF analogue, K1	471	225	21	X84798	Amino acid sequenc
399	175	15	R60133	Human G-CSF, Homo	472	226	7	P60215	Sequence encoded b
400	175	16	W08486	Platelet growth ac	473	226	9	P81029	Sequence encoded b
401	175	17	W07533	Recombinant human	474	226	13	R22673	v-sis protein p28s
402	175	17	R98465	Human granulocyte	475	229	20	X73852	Human prostate tum
403	175	17	R94394	PEGylated recombin	476	233	18	W27511	Human hypothalamic
404	175	20	X56503	Human granulocytic	477	235	20	X73974	Human prostate tum
405	175	20	W84297	Human granulocyte	478	238	20	X38577	Neisseria meningit
406	175	20	W88230	Murine Siva-1 homo	479	238	21	X74979	Neisseria meningit
407	175	21	X78936	Granulocyte colony	480	240	17	W05810	Human tumour necro
408	175	21	X78938	Granulocyte colony	481	240	17	W05810	S. pneumoniae deri
409	176	16	R79336	pMON13010 peptide.	482	241	9	P80597	CV-sis gene encode
410	176	16	R79337	pMON13037 peptide.	483	241	14	R40967	Sequence of varian
411	176	21	X53218	Human G-CSF mutant	484	241	14	R40967	CV-sis gene produc
412	176	21	X53219	Human G-CSF mutant	485	241	15	R63472	Recombinant platel
413	177	16	R86625	Human G-CSF, Homo	486	242	20	W89939	Recombinant platel
414	177	17	W02203	Human G-CSF for st	487	245	15	R47536	Antigen 3 from clu
415	177	20	W95007	Human G-CSF sequen	488	245	18	W13109	Phospholipase A2 (
416	178	8	P711031	Sequence of human	489	245	20	X13597	Human 14-3-3 zeta
417	178	13	R22646	Cancer metastasis	490	246	20	X13598	Cruciform binding
418	178	14	R32258	Human CSF, Homo s	491	246	11	R05084	Cruciform binding
419	178	16	R77378	Granulocyte colony	492	246	18	W13108	Mammalian amino ac
420	178	20	X29929	Recombinant human	493	246	19	W60199	Human 14-3-3 beta
421	178	20	X29929	T. gondii immunoge	494	246	20	X13599	Bovine 14-3-3 prot
422	179	16	R75347	Hybrid human cyto	495	248	14	R36280	Cruciform binding
423	180	12	R14443	Synthetic hg-CSF I	496	248	17	R89745	HMEL. Homo sapien
424	181	12	R13369	G-CSF deriv. witho	497	248	21	X70956	Recombinant transc
425	181	16	R87729	G-CSF circularly p	498	248	21	X28413	Arabidopsis SERK i
426	183	7	P80624	Sequence B encoded	499	249	19	W60198	14-3-3 sigma amino
427	184	16	R75344	Hybrid human cyto	500	250	17	W02278	Human mammary epit
428	185	19	W60046	Human TNF receptor	501	250	17	W29261	Anti-c-erbB-2 sing
429	185	20	X76526	Human ovarian tumo	502	250	18	W22400	Single chain antib
430	185	21	X77964	A. thaliana enviro	503	250	19	W53168	741F8 anti-c-erbB-
431	186	8	P70730	Sequence encoded b	504	250	19	W47012	Single chain bindi
432	188	19	W82717	Streptococcus pneu	505	250	20	W80422	Anti-c-erbB-2 srv'
433	189	8	P70022	Mature native gran	506	253	20	X08025	Mouse protease-rel
434	190	15	R60615	Human PDGF-B 109 s	507	253	20	W86370	Hampshire alpha me
435	190	20	X36971	Amino acid sequenc	508	253	20	W86366	Wild boar alpha me
436	191	20	X73377	Human HPDV78 prot	509	253	20	W86367	Meishan alpha mela
437	195	20	X48553	Human breast tumou	510	253	20	W86369	Large white alpha-
438	201	20	X36762	Protein involved i	511	255	20	X13596	Cruciform binding
439	204	7	P61341	Plasmid pBRV2 inse	512	258	18	W32617	Onion yellow dwarf
440	204	8	P70731	Sequence encoded b	513	260	14	R42198	BSP 14-3-3. Oryza
441	204	8	P70162	Human G-CSF encode	514	261	21	X54596	Amino acid sequenc
442	204	8	P71384	Sequence of human	515	264	20	X06930	C. albicans antige
443	204	9	P83163	Sequence of human	516	269	18	W19898	Pea type I LhcIIb
444	204	10	P95033	Granulocyte colony	517	269	21	X56860	Pea type I LhcIIb
445	204	14	R33382	Cytokine hg-CSF.	518	271	9	P80595	Recombinant platel
446	206	17	R91219	Mouse pancreas G-p	519	271	14	R40963	PDGF Bv-sis. Simi
447	206	18	W31380	Mouse MIN6-derived	520	271	15	R63468	Recombinant platel
448	206	20	W97222	MIN6-derived G pro	521	275	20	X39302	SPNF protein invol
449	206	20	W95182	Mouse MIN6-derived	522	275	21	X81826	Pig lung protease.
450	207	7	P61340	Plasmid pBRG4 gran	523	276	15	R47196	Glutamate racemase

524	6	2.0	276	18	W20575	H. pylori surface	597	317	18	W17111	Amino acid sequenc
525	6	2.0	280	19	W72134	HSV-2 strain SB5 C	598	317	18	W17113	Amino acid sequenc
526	6	2.0	281	15	R63131	Egr-1 transcriptio	599	317	18	W17105	Amino acid sequenc
527	6	2.0	282	15	R60616	Human PDGF-B 119 1	600	317	18	W21949	E6-binding protein
528	6	2.0	282	20	Y13226	Transmembrane doma	601	317	19	W77840	Multi-functional h
529	6	2.0	282	20	Y13365	Amino acid sequenc	602	317	19	W77842	Multi-functional h
530	6	2.0	282	14	R34030	FC-alpha-R. Homo	603	317	19	W77844	Multi-functional h
531	6	2.0	287	15	R99920	Human Fc-alpha-R.	604	317	19	W77846	Multi-functional h
532	6	2.0	287	20	Y33189	Human sFC-alphaRI	605	317	19	W77838	Multi-functional h
533	6	2.0	289	18	W15268	Salmonella secreta	606	319	18	W17137	Amino acid sequenc
534	6	2.0	295	21	Y68463	Mouse testis speci	607	319	18	W17139	Amino acid sequenc
535	6	2.0	296	20	Y19939	B. burgdorferi ant	608	319	19	W77857	Multi-functional h
536	6	2.0	297	20	W97733	Corn threonine syn	609	319	19	W77859	Multi-functional h
537	6	2.0	299	18	W17131	Amino acid sequenc	610	320	15	R60800	Rape abscission/de
538	6	2.0	299	20	W89198	Aspartate-specific	611	320	18	W17115	Amino acid sequenc
539	6	2.0	300	19	W82558	C. elegans glycoge	612	320	18	W17117	Amino acid sequenc
540	6	2.0	300	20	Y13941	Human transmembran	613	320	18	W17119	Amino acid sequenc
541	6	2.0	302	18	W17106	Amino acid sequenc	614	320	18	W17121	Amino acid sequenc
542	6	2.0	302	18	W17108	Amino acid sequenc	615	320	18	W17123	Amino acid sequenc
543	6	2.0	302	18	W17110	Amino acid sequenc	616	320	19	W77856	Multi-functional h
544	6	2.0	302	18	W17112	Amino acid sequenc	617	320	19	W77848	Multi-functional h
545	6	2.0	302	18	W17104	Amino acid sequenc	618	320	19	W77850	Multi-functional h
546	6	2.0	302	19	W77841	Multi-functional h	619	320	19	W77852	Multi-functional h
547	6	2.0	302	19	W77843	Multi-functional h	620	320	19	W77854	Multi-functional h
548	6	2.0	302	19	W77845	Multi-functional h	621	321	18	W17124	Amino acid sequenc
549	6	2.0	302	19	W77837	Multi-functional h	622	321	18	W17125	Amino acid sequenc
550	6	2.0	302	19	W77839	Multi-functional h	623	321	18	W17126	Amino acid sequenc
551	6	2.0	302	21	Y50904	Human endometriosi	624	321	18	W17127	Amino acid sequenc
552	6	2.0	303	20	Y73636	S. pneumoniae lacC	625	321	19	W77832	Multi-functional h
553	6	2.0	304	20	Y09294	Rice beta-glucanas	626	321	19	W77829	Multi-functional h
554	6	2.0	305	18	W17114	Amino acid sequenc	627	321	19	W77830	Multi-functional h
555	6	2.0	305	18	W17116	Amino acid sequenc	628	321	19	W77831	Multi-functional h
556	6	2.0	305	18	W17118	Amino acid sequenc	629	322	16	R79330	IL-3 containing fu
557	6	2.0	305	18	W17120	Amino acid sequenc	630	322	16	R79332	IL-3 containing fu
558	6	2.0	305	18	W17122	Amino acid sequenc	631	322	16	R79333	IL-3 containing fu
559	6	2.0	305	19	W77849	Multi-functional h	632	322	16	R79305	IL-3 containing fu
560	6	2.0	305	19	W77851 ^h	Multi-functional h	633	322	16	R79306	IL-3 containing fu
561	6	2.0	305	19	W77853	Multi-functional h	634	322	16	R79307	IL-3 containing fu
562	6	2.0	305	19	W77855	Multi-functional h	635	322	16	R79315	IL-3 containing fu
563	6	2.0	305	19	W77847	Multi-functional h	636	322	16	R79325	IL-3 containing fu
564	6	2.0	307	16	R79328	IL-3 containing fu	637	322	16	R79326	IL-3 containing fu
565	6	2.0	307	16	R79334	IL-3 containing fu	638	322	18	W17138	Amino acid sequenc
566	6	2.0	307	16	R79335	IL-3 containing fu	639	322	18	W17140	Amino acid sequenc
567	6	2.0	307	16	R79298	IL-3 containing fu	640	322	19	W77858	Multi-functional h
568	6	2.0	307	16	R79299	IL-3 containing fu	641	322	19	W77860	Multi-functional h
569	6	2.0	307	16	R79300	IL-3 containing fu	642	322	21	Y53186	Human interleukin-
570	6	2.0	307	16	R79301	IL-3 containing fu	643	322	21	Y53187	Human interleukin-
571	6	2.0	307	16	R79311	IL-3 containing fu	644	322	21	Y53188	Human interleukin-
572	6	2.0	307	16	R79312	IL-3 containing fu	645	322	21	Y53196	Human interleukin-
573	6	2.0	307	16	R79322	IL-3 containing fu	646	322	21	Y53206	Human interleukin-
574	6	2.0	307	16	R79323	IL-3 containing fu	647	322	21	Y53207	Human interleukin-
575	6	2.0	307	21	Y53179	Human interleukin-	648	322	21	Y53211	Human interleukin-
576	6	2.0	307	21	Y53180	Human interleukin-	649	322	21	Y53213	Human interleukin-
577	6	2.0	307	21	Y53181	Human interleukin-	650	322	21	Y53214	Human interleukin-
578	6	2.0	307	21	Y53182	Human interleukin-	651	323	20	Y29153	Amino acid sequenc
579	6	2.0	307	21	Y53192	Human interleukin-	652	323	20	Y19938	B. burgdorferi ant
580	6	2.0	307	21	Y53193	Human interleukin-	653	326	19	W78110	Flt3l/GS/G-CSF pol
581	6	2.0	307	21	Y53203	Human interleukin-	654	327	14	R41688	Murine Fas. Mus m
582	6	2.0	307	21	Y53204	Human interleukin-	655	327	16	R78611	Murine Fas antigen
583	6	2.0	307	21	Y53209	Human interleukin-	656	327	17	R92530	mFas sequence. Sy
584	6	2.0	307	21	Y53215	Human interleukin-	657	327	20	W86241	Fas ligand (FasL)
585	6	2.0	307	21	Y53216	Human interleukin-	658	329	18	W17128	Amino acid sequenc
586	6	2.0	308	19	W60946	Streptococcus pneu	659	329	18	W17129	Amino acid sequenc
587	6	2.0	308	20	Y28816	pm4_l3 secreted pr	660	329	18	W17130	Amino acid sequenc
588	6	2.0	308	20	Y08454	Yeast Padl protein	661	329	18	W17132	Amino acid sequenc
589	6	2.0	309	13	R27754	Norwalk virus 8FII	662	329	18	W17133	Amino acid sequenc
590	6	2.0	309	17	R98528	Cyclase encoded by	663	329	18	W17134	Amino acid sequenc
591	6	2.0	310	17	W04931	Cellulytic enzyme	664	329	18	W17135	Amino acid sequenc
592	6	2.0	313	13	R27753	Norwalk virus 8FII	665	329	18	W17136	Amino acid sequenc
593	6	2.0	314	20	R31222	L. vulgaris DFPase	666	329	19	W77898	Multi-functional h
594	6	2.0	317	13	R23600	Recombinant hemato	667	329	19	W77899	Multi-functional h
595	6	2.0	317	18	W17107	Amino acid sequenc	668	329	19	W77900	Multi-functional h
596	6	2.0	317	18	W17109	Amino acid sequenc	669	329	19	W77901	Multi-functional h

670	6	2.0	329	19	w77902	Multi-functional h	743	6	2.0	357	16	R85085	Human glucose-6-ph
671	6	2.0	329	19	w77833	Multi-functional h	744	6	2.0	357	16	R85086	Human glucose-6-ph
672	6	2.0	329	19	w77834	Multi-functional h	745	6	2.0	357	16	R85088	Murine wild type g
673	6	2.0	329	19	w77835	Multi-functional h	746	6	2.0	357	16	R85083	Human wild type g1
675	6	2.0	329	19	w77836	Multi-functional h	747	6	2.0	359	19	w78076	Flt3L (39/40 break
676	6	2.0	330	16	w79969	Geranylgeranyl dip	748	6	2.0	360	20	w97621	Human neuregulin r
677	6	2.0	330	18	w12386	Geranylgeranyl dip	749	6	2.0	361	17	w00378	Fusion peptide #2
678	6	2.0	330	18	w12387	Geranylgeranyl dip	750	6	2.0	361	18	w34336	Tomato mottle viru
679	6	2.0	330	18	w12388	Geranylgeranyl dip	751	6	2.0	361	18	w34324	Tomato mottle viru
680	6	2.0	330	18	w12389	Geranylgeranyl dip	752	6	2.0	361	18	w34325	Tomato mottle viru
681	6	2.0	330	18	w12390	Geranylgeranyl dip	753	6	2.0	361	18	w34326	Tomato mottle viru
682	6	2.0	330	18	w12391	Geranylgeranyl dip	754	6	2.0	363	19	w78074	Flt3L/IgG2b/G-CSF
683	6	2.0	330	19	w62529	Mutant geranylgera	755	6	2.0	364	19	w78095	Flt3L (89/90 break
684	6	2.0	330	19	w62530	Mutant geranylgera	756	6	2.0	364	19	w78079	Flt3L (34/35 break
685	6	2.0	330	19	w62531	Mutant geranylgera	757	6	2.0	364	19	w78084	Flt3L (65/66 break
686	6	2.0	330	19	w62528	Geranylgeranyl dip	758	6	2.0	364	19	w78085	Flt3L (98/99 break
687	6	2.0	330	19	w52282	Geranylgeranyl dip	759	6	2.0	364	19	w78090	Flt3L (65/66 break
688	6	2.0	330	19	w52283	Mutant prenilyl dip	760	6	2.0	364	19	w78093	G-CSF/IgG2b/Flt3L
689	6	2.0	331	20	w09302	Rice Gns5 CBS prot	761	6	2.0	364	19	w78061	G-CSF/IgG2b/Flt3L
690	6	2.0	331	20	w67889	Human secreted pro	762	6	2.0	364	19	w78077	Flt3L (39/40 break
691	6	2.0	332	21	w52284	Mutant prenilyl dip	763	6	2.0	365	17	w00379	Fusion peptide #3
692	6	2.0	332	21	w92569	Rhodothermus obame	764	6	2.0	365	20	w89557	Triticum sp. cyste
693	6	2.0	334	15	r47241	Wild-type Feline H	765	6	2.0	366	19	w78092	Flt3L/IgG2b/G-CSF
694	6	2.0	334	19	w83290	Human flt3 ligand	766	6	2.0	367	19	w78114	Sequence rearrange
695	6	2.0	334	19	w78111	Flt3L/GS/G-CSF pol	767	6	2.0	369	20	y16629	Protein encoded b
696	6	2.0	334	19	w78112	Flt3L/GS/G-CSF pol	768	6	2.0	369	20	y07080	Renal cancer assoc
697	6	2.0	334	19	w78108	Flt3L/GS/G-CSF pol	769	6	2.0	370	17	r91218	Human pituitary G-
698	6	2.0	334	19	w78109	Flt3L/GS/G-CSF pol	770	6	2.0	370	18	w31379	Human G protein-co
699	6	2.0	334	19	w77821	G-CSF-Flt-3 ligand	771	6	2.0	370	19	w78096	Flt3L/IgG2b/G-CSF
700	6	2.0	335	16	r79319	IL-3 containing fu	772	6	2.0	370	19	w78099	Flt3L/IgG2b/G-CSF
701	6	2.0	335	13	y53200	Human interleukin-	773	6	2.0	370	19	w78105	Chimeric G-CSF-Flt
702	6	2.0	336	13	r23077	IL-3:G-CSF, recomb	774	6	2.0	370	19	w78107	Chimeric G-CSF-Flt
703	6	2.0	337	16	r79324	IL-3 containing fu	775	6	2.0	370	19	w78078	Flt3L (39/40 break
704	6	2.0	337	21	y53205	Human interleukin-	776	6	2.0	370	19	w78080	Flt3L (35/36 break
705	6	2.0	339	19	w60076	Escherichia coli R	777	6	2.0	370	19	w78081	Flt3L (36/37 break
706	6	2.0	340	20	y42127	Rice tryptophan sy	778	6	2.0	370	19	w78082	Flt3L (37/38 break
707	6	2.0	340	21	y84360	Amino acid sequenc	779	6	2.0	370	19	w78083	Flt3L/IgG2b/G-CSF
708	6	2.0	342	15	y58681	Wheat N-acetylglut	780	6	2.0	370	19	w78086	Flt3L/IgG2b/G-CSF
709	6	2.0	342	17	w02660	G-protein coupled	781	6	2.0	370	19	w78087	Flt3L/IgG2b/G-CSF
710	6	2.0	342	13	r22549	G-protein coupled	782	6	2.0	370	19	w78088	Flt3L/IgG2b/G-CSF
711	6	2.0	344	17	r31439	Human macrophage-s	783	6	2.0	370	19	w78089	Flt3L/IgG2b/G-CSF
712	6	2.0	344	19	w83304	Human FCRI (CDNA c	784	6	2.0	370	19	w78091	Flt3L (S15F)/IgG2b
713	6	2.0	345	20	y37548	Human flt3 ligand	785	6	2.0	370	19	w78055	G-CSF/IgG2b/Flt3L
714	6	2.0	346	16	R85087	Amino acid sequenc	786	6	2.0	370	19	w78056	G-CSF/IgG2b/Flt3L
715	6	2.0	346	16	w23088	Human glucose-6-ph	787	6	2.0	370	19	w78057	G-CSF/IgG2b/Flt3L
716	6	2.0	346	21	w82743	Pyrodictium sp. es	788	6	2.0	370	19	w78058	G-CSF/IgG2b/Flt3L
717	6	2.0	346	21	w82743	DNA replication an	789	6	2.0	370	19	w78059	G-CSF/IgG2b/Flt3L
718	6	2.0	347	19	w78100	Flt3L/IgG2b/G-CSF	790	6	2.0	370	19	w78060	G-CSF/IgG2b/Flt3L
719	6	2.0	347	19	w78103	Chimeric receptor	791	6	2.0	370	19	w78063	G-CSF/IgG2b/Flt3L
720	6	2.0	348	19	w78075	Flt3L/IgG2b/G-CSF	792	6	2.0	370	19	w78071	Flt3L/IgG2b/G-CSF
721	6	2.0	348	20	y25806	Human secreted pro	793	6	2.0	370	19	w78072	Flt3L/IgG2b/G-CSF
722	6	2.0	348	21	y53612	Amino acid sequenc	794	6	2.0	370	19	w78054	G-CSF/IgG2b/Flt3L
723	6	2.0	349	16	r79316	IL-3 containing fu	795	6	2.0	370	19	w78822	G-CSF-linker-Flt3L
724	6	2.0	349	16	r79316	IL-3 containing fu	796	6	2.0	370	20	w97221	Human pituitary-de
725	6	2.0	349	17	w00377	Fusion peptide #1	797	6	2.0	370	20	w95181	Human G-protein co
726	6	2.0	349	18	w12719	PhD gene product.	798	6	2.0	370	21	y87503	Murine G coupled-p
727	6	2.0	349	19	w83286	Human flt3 ligand	799	6	2.0	370	21	y87505	Human G coupled-pr
728	6	2.0	349	19	w83287	Human flt3 ligand	800	6	2.0	370	21	y87507	Rat G coupled-prot
729	6	2.0	349	19	w83288	Human flt3 ligand	801	6	2.0	374	19	w97834	Human Fc receptor
730	6	2.0	349	19	w83289	Human flt3 ligand	802	6	2.0	375	18	w15512	Human MAP kinase #1
731	6	2.0	349	19	w83290	Chimeric receptor	803	6	2.0	375	19	w59176	MAP kinase #1. Sy
732	6	2.0	349	19	w78003	G-CSF/IgG2b/Flt3L	804	6	2.0	376	20	y37222	P. calinivirus act
733	6	2.0	349	19	w78005	Flt3L 1-139/IgG2b/	805	6	2.0	376	20	y37222	Protein which is s
734	6	2.0	349	20	y16578	A 2,4-diacetylphlo	806	6	2.0	378	19	w78094	Flt3L/IgG2b/G-CSF
735	6	2.0	349	21	y84617	Amino acid sequenc	807	6	2.0	378	19	w78097	Flt3L/IgG2b/G-CSF
736	6	2.0	349	21	y53197	Human interleukin-	808	6	2.0	378	19	w78098	G-CSF/IgG2b/G-CSF
737	6	2.0	349	21	y53208	Human interleukin-	809	6	2.0	378	19	w78062	G-CSF/IgG2b/Flt3L
738	6	2.0	352	19	w78113	Sequence rearrange	810	6	2.0	378	20	y41682	Murine D6 protein.
739	6	2.0	352	19	w78073	Flt3L/IgG2b/G-CSF	811	6	2.0	380	18	w27510	Consensus human hy
740	6	2.0	353	19	w78104	Chimeric Flt3L-G-C	812	6	2.0	380	20	y29326	Human secreted pro
741	6	2.0	353	19	w78106	Chimeric Flt3L-G-C	813	6	2.0	382	19	w76190	Actinoplanes sp. a
742	6	2.0	355	19	w77820	Flt3L-linker-G-CSF	814	6	2.0	382	21	y74705	Neisseria gonorrhe
	6	2.0	355	21	y84817	Amino acid sequenc	815	6	2.0	382	21	y74706	Neisseria meningit

816	6	2.0	384	19	W78101	Chimeric receptor	889	20	Y28604	L. cuprina EcR par			
817	6	2.0	386	20	W98200	RTD, inhibitor of	890	2.0	459	10	P96202	Human muscarinic a	
818	6	2.0	386	20	Y04144	Human Tango-74 pro	891	6	2.0	463	21	Y52216	Arabidopsis thalia
819	6	2.0	386	20	W99018	Human TRAIL recept	892	6	2.0	465	20	Y31642	Human transport-as
820	6	2.0	386	20	W99019	Human TRAIL recept	893	6	2.0	468	20	W94850	Mouse MAPK kinase
821	6	2.0	386	20	W92792	Human TNF receptor	894	6	2.0	468	20	W94853	Mouse mutant MAPK
822	6	2.0	386	21	Y69991	Human receptor-ass	895	6	2.0	471	6	P50308	Cellobiohydrolase
823	6	2.0	389	20	W97663	Mouse mitogen acti	896	6	2.0	471	16	R77262	T. longibrachiatum
824	6	2.0	392	18	W09039	Mitogen-activated	897	6	2.0	471	17	W02025	Trichoderma cellob
825	6	2.0	398	12	R10572	B.subtilis rib ORF	898	6	2.0	471	19	W61216	Streptococcus pneu
826	6	2.0	398	12	R15614	Human type II inte	899	6	2.0	472	18	W23598	Human LYST2 polype
827	6	2.0	398	14	R42061	Lymphoblastoid der	900	6	2.0	472	20	Y32120	Human LYST-2 prote
828	6	2.0	398	15	R60617	Human type II IL-1	901	6	2.0	473	13	R21977	RuBPCase large sub
829	6	2.0	398	16	R85480	Human type II inte	902	6	2.0	475	17	Y05777	Human peroxisome p
830	6	2.0	398	18	W24095	L-methionine gamma	903	6	2.0	475	17	R99327	Peroxisome prolif
831	6	2.0	398	18	W10236	Pseudomonas putida	904	6	2.0	475	17	R99328	Peroxisome prolif
832	6	2.0	398	19	W59267	Human Type II IL-1	905	6	2.0	475	17	R92478	Peroxisome prolif
833	6	2.0	398	20	Y55941	Human PAK5 protein	906	6	2.0	475	20	W99596	Mouse peroxisome p
834	6	2.0	398	20	W96261	P. putida methioni	907	6	2.0	477	17	R99324	Peroxisome prolif
835	6	2.0	398	20	W93426	P. putida methioni	908	6	2.0	477	17	R99327	Peroxisome prolif
836	6	2.0	398	20	Y01096	Methioninase proti	909	6	2.0	478	20	Y14051	G. oxydans D-sorbi
837	6	2.0	398	20	W81552	Bacillus subtilis	910	6	2.0	478	20	Y05471	Human PPAR-gammal
838	6	2.0	399	20	Y22573	Bacterial general	911	6	2.0	479	15	R60695	Fragment of the mu
839	6	2.0	401	20	Y03143	Heparin sulphate 6	912	6	2.0	480	11	R06240	Soluble intercellu
840	6	2.0	401	20	Y03144	Heparin sulphate 6	913	6	2.0	481	17	W16839	Recombinant endoto
841	6	2.0	401	21	Y96284	Sinorhizobium melli	914	6	2.0	481	17	W19010	Recombinant endoto
842	6	2.0	401	21	Y83903	Mouse HS6S73 prote	915	6	2.0	481	17	W19011	Recombinant endoto
843	6	2.0	410	19	W68414	Mycobacterium bovi	916	6	2.0	483	21	Y91433	Human secreted pro
844	6	2.0	410	20	Y03140	Heparin sulphate 6	917	6	2.0	487	17	W16817	Recombinant endoto
845	6	2.0	410	21	Y83900	Human HS6S71 prote	918	6	2.0	487	21	Y75687	Neisseria gonorrh
846	6	2.0	410	21	Y25290	Human prostate gro	919	6	2.0	487	21	Y75688	Neisseria meningit
847	6	2.0	411	20	W83358	Mycobacterium tube	920	6	2.0	487	21	Y75689	Neisseria meningit
848	6	2.0	414	17	Y10529	Saccharomyces cere	921	6	2.0	489	14	R42176	Murine MDM2. Mus
849	6	2.0	415	21	Y74283	Neisseria gonorrh	922	6	2.0	489	16	R76697	Mouse MDM2 protein
850	6	2.0	415	21	Y74286	Neisseria gonorrh	923	6	2.0	489	17	W07888	Murine MDM-2, invo
851	6	2.0	417	13	R22095	Phosphoglycerate k	924	6	2.0	489	18	W15464	Murine MDM2. Mus
852	6	2.0	417	19	W54355	47 kD heat shock p	925	6	2.0	489	19	W57246	Mouse MDM2 protein
853	6	2.0	417	21	Y81908	Human Hsp47 protei	926	6	2.0	489	19	W48242	Mouse MDM2. Mus m
854	6	2.0	418	13	R22025	A. chrysogenum pho	927	6	2.0	489	19	W42997	Amino acid sequenc
855	6	2.0	418	14	R49247	PGK. Acremonium c	928	6	2.0	489	19	W42972	Amino acid sequenc
856	6	2.0	422	20	Y00251	Enterococcus faeca	929	6	2.0	489	20	W94305	Mouse MDM2. Mus m
857	6	2.0	423	18	W15513	MAP kinase #2. Sy	930	6	2.0	490	21	Y81745	Streptococcus pneu
858	6	2.0	425	20	W98019	Mouse calcium acti	931	6	2.0	491	13	R27787	Adrenodoxin reduct
859	6	2.0	428	21	Y94897	Human protein clon	932	6	2.0	494	17	R99323	Peroxisome prolif
860	6	2.0	430	21	Y44388	Zea mays RecA-like	933	6	2.0	496	20	W76984	Mouse Egr-1 protei
861	6	2.0	431	19	W76434	Human p53 regulate	934	6	2.0	497	13	R27786	Adrenodoxin reduct
862	6	2.0	431	20	Y36903	Protein involved i	935	6	2.0	497	16	R66693	Human adrenodoxin-
863	6	2.0	432	16	R77865	S. clavuligerus OR	936	6	2.0	497	20	W99362	C.trachomatis Cyss
864	6	2.0	433	12	R13838	Fusaric acid resis	937	6	2.0	497	20	W95461	L. helveticus pept
865	6	2.0	433	19	W41732	Arabidopsis chloro	938	6	2.0	497	21	Y74865	Neisseria meningit
866	6	2.0	433	21	Y80093	Permease foci amin	939	6	2.0	497	21	Y75810	Neisseria meningit
867	6	2.0	434	19	W50008	Human hyaluronidas	940	6	2.0	497	21	Y75882	Neisseria strain 2
868	6	2.0	434	20	Y15215	Amino acid sequenc	941	6	2.0	498	15	R55800	Interleukin 14. H
869	6	2.0	435	20	W94849	Human MAPK kinase	942	6	2.0	499	20	Y34122	Human potassium ch
870	6	2.0	436	12	R14533	Truncated Protein	943	6	2.0	500	20	Y37360	Protein involved 1
871	6	2.0	437	20	W67722	Human tumour antiq	944	6	2.0	505	17	R99326	Peroxisome prolif
872	6	2.0	437	21	Y87261	Human signal pepti	945	6	2.0	508	18	W14721	Human ICAM-1 (del4
873	6	2.0	437	21	Y66737	Membrane-bound pro	946	6	2.0	509	19	W47533	Amino acid sequenc
874	6	2.0	438	19	W59836	Maize Id protein.	947	6	2.0	509	19	W26792	Mouse TIE-2 recept
875	6	2.0	438	21	Y55985	S.clavuligerus cla	948	6	2.0	509	20	Y23734	TIE ligand-3 (TL3)
876	6	2.0	441	15	R55826	GDF-9. Homo sapi	949	6	2.0	510	20	Y29618	Human PIAS1 protei
877	6	2.0	441	20	Y09152	Growth differentia	950	6	2.0	514	20	W84268	Candida albicans T
878	6	2.0	447	18	W13381	Sesame omega-3 ali	951	6	2.0	515	15	R48670	Chitinase derivati
879	6	2.0	448	19	W72048	HSV-2 strain SB5 C	952	6	2.0	518	20	W97737	Soybean threonine
880	6	2.0	448	21	Y44840	Human orphan cyto	953	6	2.0	523	19	W78006	Trimeric Flt3L-G-C
881	6	2.0	449	19	W79140	Plant acetolactate	954	6	2.0	523	19	W78008	Trimeric Flt3L-G-C
882	6	2.0	449	20	Y35696	C. pneumoniae prot	955	6	2.0	528	18	W20908	H. pylori inner me
883	6	2.0	449	20	Y00250	Enterococcus faeca	956	6	2.0	530	18	W28501	Birch pollen co-fa
884	6	2.0	450	16	R70030	UTH1 gene product.	957	6	2.0	531	16	R80110	ICAM-1 CD54. Homo
885	6	2.0	450	19	W46606	Tyrosine kinase as	958	6	2.0	532	10	P91357	Intercellular adhe
886	6	2.0	450	20	Y29635	Human tyrosine kin	959	6	2.0	532	11	R04165	Intercellular adh
887	6	2.0	450	20	Y22130	Human tyrosine kin	960	6	2.0	532	13	R20809	Intercellular adh
888	6	2.0	456	19	W72032	HSV-2 strain SB5 C	961	6	2.0	532	14	R35071	ICAM-1. Homo sapi

962	6	2.0	532	15	R46066	Human ICAM-1. Hom
963	6	2.0	532	15	R58779	Inter-cellular adh
964	6	2.0	532	16	R79457	ICAM-1. Homo sapi
965	6	2.0	532	17	R91437	Human ICAM-1. Hom
966	6	2.0	532	17	R90294	Intracellular adhe
967	6	2.0	532	18	R72720	Human intracellular
968	6	2.0	532	18	W14720	Human ICAM-1. Hom
969	6	2.0	532	18	W09313	Human ICAM-1 (enco
970	6	2.0	532	19	W80446	Human intracellular
971	6	2.0	532	19	W70871	Intracellular adhe
972	6	2.0	532	19	W71263	Human intercellula
973	6	2.0	532	19	W46735	Amino acid sequenc
974	6	2.0	532	20	W86193	Human intracellular
975	6	2.0	532	21	Y59499	Human ICAM-1 prote
976	6	2.0	532	10	P93113	Egr-1. Mus. p931
977	6	2.0	533	15	R63129	Mouse Egr-1 clone
978	6	2.0	536	19	W81732	M. tuberculosis im
979	6	2.0	536	19	W43365	Mycobacterium tube
980	6	2.0	536	20	Y39162	M. tuberculosis an
981	6	2.0	536	20	Y39019	M. tuberculosis re
982	6	2.0	536	21	Y52525	House dust mite (D
983	6	2.0	541	19	W44165	Entodinium caudatu
984	6	2.0	542	20	Y26973	Arabidopsis thalia
985	6	2.0	543	19	W72196	HSV-2 strain SB5 C
986	6	2.0	544	19	W78118	Trimeric Flt3L-G-C
987	6	2.0	550	19	W78117	Trimeric Flt3L-G-C
988	6	2.0	554	16	R80444	Human soluble epox
989	6	2.0	555	19	W70480	Sindbis virus nsp3
990	6	2.0	555	21	Y52523	House dust mite (D
991	6	2.0	556	21	Y73342	Human signal pepti
992	6	2.0	557	18	W28504	Mugwort grass pol
993	6	2.0	562	10	P80458	Sequence of human
994	6	2.0	562	19	W78121	Trimeric Flt3L-G-C
995	6	2.0	565	19	W78119	Trimeric Flt3L-G-C
996	6	2.0	565	20	W96316	Acidic leucine ami
997	6	2.0	567	20	Y28998	Human TGF-beta act
998	6	2.0	567	21	Y81722	Streptococcus pneu
999	6	2.0	568	16	R66452	Serine-rich AF-9 p
1000	6	2.0	571	20	W96315	Acidic leucine ami

ALIGNMENTS

RESULT 1
W6102 ID W66102 standard; Protein: 300 AA.

XX AC W66102;

XX DT 02-DEC-1998 (first entry)

XX DE Amino acid sequence of tumour necrosis related receptor (TR4).

XX KW Human; tumour necrosis related receptor; TR4; agonist; antagonist;
KW inhibition; chronic; acute; inflammation; arthritis; septicaemia;
KW autoimmune disease; transplant rejection; stroke; cancer;
XX Alzheimer's disease.

XX OS Homo sapiens.

XX PN EP861850-A1.

XX PD 02-SEP-1998.

XX PF 20-JAN-1998; 98EP-0300382.

XX PR 04-FEB-1997; 97US-0794796.

XX PA (SMIK) SMITHKLINE BEECHAM CORP.

XX PI Emery J, Tan KB, Truneh A, Young PR;

XX XX

DR WPI: 1998-508248/44.
XX N-PSDB; V07654.
XX New DNA encoding tumour necrosis related receptor - used to treat
PT and prevent e.g. inflammation, arthritis, septicaemia, autoimmune
PT diseases, transplant rejection, infection, stroke, ischaemia, ARDS,
PT restenosis, AIDS, bone disorders and cancer
XX
XX Claim 1; Fig 1; 21pp; English.
XX
CC This is the amino acid sequence of the human tumour necrosis related
CC receptor (TR4), used in the method of the invention. The TR4 protein
CC or its agonist can be used to treat a subject in need of enhanced
CC TR4 polypeptide activity. The antagonist is used to inhibit TR4
CC polypeptide activity. The active agents can be used for the
CC treatment and prevention of diseases such as chronic and acute
CC inflammation, arthritis, septicaemia, autoimmune diseases, transplant
CC rejection, stroke, cancer, Alzheimer's disease.
XX
XX Sequence 300 AA;

Query Match 100.0%; Score 300; DB 19; Length 300;
Best Local Similarity 100.0%; Pred. No. 5.6e-270; Indels 0; Gaps 0;
Matches 300; Conservative 0; Mismatches 0;

QY 1 MRALEGGSLCLCLVLPALLPVPVAVRGVYAEPTTPWRDAETGERLVCAQCPTGTFVQR 60
|||||
Db 1 mraleggslclclvlpallpvpavrgvyaetpypwrdaetgerlvcaqcptgtfvqr 60
QY 61 PCRRDPTTCGCPGPPRHYYQFQWYLERCRYCNVLCGEREEARACHATHNRACRGTGFF 120
|||||
Db 61 pcrrdpttcgcpgrprrhytqfwnylercrycnvlgcgeraachathnracrtgff 120
QY 121 AHAGFCLEHASCPGAGVIAAGTPSONTOCCPCPGTFESASSSSSEQCOPHRNCTALGLA 180
|||||
Db 121 ahagfclenascpaggavtaptgpcpntgcpcpgtfesassssseqcpnrcnctalgla 180
QY 181 LNVPGSSSHDTLCTCTGTFPLSTRVPGAECERAVIDFAFQDISIKRLQLQALEAPE 240
|||||
Db 181 lnvpgssshdtlctctgtgfpstrvpgaecercaravidfafqdisikrlqlqaleape 240
QY 241 GWGPTPRAGAAQLKLRRLTELLGAQDQGALLVRLQLALVARMPGLERSVRERFLPVH 300
|||||
Db 241 gwgptpragaaqlklrrrltellgaqdqgallvrlqlalvarmpglersvrerflpvh 300

RESULT 2

W63622 ID W63622 standard; Protein: 300 AA.

XX AC W63622;

XX DT 26-OCT-1998 (first entry)

XX DE Human tumour necrosis factor receptor-6 alpha protein.

XX KW Human tumour necrosis factor receptor-6 alpha; TNFR-6 alpha; TNFR-6 beta;
KW endothelial cells; keratinocytes; normal prostate; apoptosis;
KW prostate tumour tissue.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Peptide 1..30

FT Protein 31..300 /note= "TNFR-6 alpha"

FT Region 31..282 /note= "Soluble extracellular domain"

XX PN WO9830694-A2.

XX XX

XX PN 16-JUL-1998.

XX PD

XX PF 13-JAN-1998; 98WO-US00153.
XX PR 14-JAN-1997; 97US-0035496.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Ebner R, Feng P, Gentz RL, Ni J, Ruben SM, Yu G;
XX DR WPI; 1998-399142/34.
XX DR N-PSDB; V39085.
XX PT Human tumour necrosis factor receptors 6-alpha and 6-beta - used in
XX PT the diagnosis of immune system-related disorder(s)
XX PS Claim 20; Fig 1; 91pp; English.
XX CC The present sequence represents the human tumour necrosis factor
XX CC receptor-6 alpha (TNFR-6 alpha) protein. The invention also provides
XX CC for the TNFR-6 beta protein (W63623). TNFR-6 alpha and TNFR-6 beta are
XX CC members of the tumour necrosis factor receptor (TNFR) family. TNFRs
XX CC are expressed in endothelial cells, keratinocytes, normal prostate and
XX CC prostate tumour tissue. For a number of disorders of these cells,
XX CC increased or decreased levels of TNFR-6 alpha and/or TNFR-6 beta
XX CC expression can be detected, therefore the TNFR-6 alpha and TNFR-6 beta
XX CC polypeptides, nucleic acids and antibodies are claimed to be useful in
XX CC the diagnosis of such disorders. Mutations of the TNFR-6 alpha and
XX CC TNFR-6 beta genes can also be detected. The TNFR polypeptides are
XX CC also claimed to be useful for identifying ligands which may be useful
XX CC in the treatment of apoptosis related disorders.
XX SQ Sequence 300 AA;
Query Match 100.0%; Score 300; DB 19; Length 300;
Best Local Similarity 100.0%; Pred. No. 5.6e-270;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MRALEGGLSLCLVLPALLPVPARGVAETPTYPWRDAETGERLVCAQCPGTFVQR 60
Db 1 mralegpglsllclvlpallpvpargvaetptyprwdaetgerlvcaqcpptfvqr 60
Qy 61 PCRDSPTTCGCPPRHYTQFWNYLRCRYCNVLCGEREEARACHATHNACRCRTGFF 120
Db 61 pcrdspttcgcpprhytqfwnylrcrycnvlgereearachathnacrctgff 120
Qy 121 AHAGFCLEHASCPPGAGVIAPGTPSQNTQCPCPGTFSASSSSSEQCPHRNCTALGLA 180
Db 121 ahagfclehascpaggviapgtpsqntqcpcpgtfsassssseqcqhprnctalgla 180
Qy 181 LNVPGSSSHDILCTSCGTFPLSTRVPGAEECEERAVIDFVAFQDISIKRLQLQALEAPE 240
Db 181 lnpvgssshdiltctscgtfplstrvpgaeeceeravidfvafqdisikrlqlqaleape 240
Qy 241 GWGPTPRAGRAALQKLRRRLTELLGAGDGLLVRLQALVARMPGLERSVRERFLPVH 300
Db 241 gwgptpragraalqlkrrrltellgagdgallvrlqalvarmpglersvrerflpvh 300
RESULT 3
Y03099
ID Y03099 standard; Protein; 300 AA.
XX AC Y03099;
XX XX
XX 09-DEC-1999 (first entry)
XX DE Human lung TNF-receptor protein.
XX XX
XX Tumour necrosis factor; TNF; TNF receptor; human; lung; gene therapy;
KW detection; immunoassay; diagnosis; disease; immune system; tumour;
KW osteogenic system; cardiovascular system; central nervous system; asthma;

KW peripherical nervous systems; transplant incompatibility; antitumor;
KW rheumatoid arthritis; antiasthmatic; antiarthritic.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 134..1036
FT /*tag= a
FT /product= "TNF-receptor"
XX PN DE19809978-AL.
XX PD 16-SEP-1999.
XX PF 09-MAR-1998; 98DE-1009978.
XX PR 09-MAR-1998; 98DE-1009978.
XX PA (BADI) BASF AG.
XX PI Kroeger B;
XX DR WPI; 1999-519473/44.
XX DR N-PSDB; 209998.
XX PT New soluble member of tumor necrosis factor receptor family, useful for
XX PT identification specific modulators and for treating disease e.g. tumors
XX PS Claim 1; Page 8-9; 10pp; German.
XX CC This invention describes a novel tumour necrosis factor (TNF) receptor
XX CC (I) isolated from human lung tissue. (I) is used: (i) to raise specific
XX CC antibodies (Ab); (ii) to screen for specific (therapeutic) ligands
XX CC (A), potential therapeutic agents; and (iii) therapeutically (optionally
XX CC expressed from a gene therapy vector) in conditions associated with a
XX CC deficit of (I). Ab are used: (a) for qualitative or quantitative
XX CC detection of (I) in standard immunoassays (for diagnosis of disease, or
XX CC susceptibility, or for monitoring); and (b) as therapeutic inhibitors in
XX CC cases where (I) is overexpressed. Nucleic acid (II) that encodes (I) is
XX CC used: (A) for recombinant production of (I); (B) also its oligonucleotide
XX CC fragments, in standard hybridization and/or amplification assays; (C) as
XX CC source of antisense molecules or ribozymes; and (D) to produce transgenic
XX CC animals (for studying (patho)physiology of (I)). Diseases possibly
XX CC associated with under- or over-expression of (I) are those of the immune,
XX CC osteogenic, cardiovascular and central or peripheral nervous systems,
XX CC tumors, transplant incompatibility, asthma and rheumatoid arthritis. The
XX CC products of the invention have antitumor, antiasthmatic and
XX CC antiarthritic activity. This sequence represents the TNF-receptor of the
XX CC invention.
XX SQ Sequence 300 AA;
Query Match 100.0%; Score 300; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 5.6e-270;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MRALEGGLSLCLVLPALLPVPARGVAETPTYPWRDAETGERLVCAQCPGTFVQR 60
Db 1 mralegpglsllclvlpallpvpargvaetptyprwdaetgerlvcaqcpptfvqr 60
Qy 61 PCRDSPTTCGCPPRHYTQFWNYLRCRYCNVLCGEREEARACHATHNACRCRTGFF 120
Db 61 pcrdspttcgcpprhytqfwnylrcrycnvlgereearachathnacrctgff 120
Qy 121 AHAGFCLEHASCPPGAGVIAPGTPSQNTQCPCPGTFSASSSSSEQCPHRNCTALGLA 180
Db 121 ahagfclehascpaggviapgtpsqntqcpcpgtfsassssseqcqhprnctalgla 180
Qy 181 LNVPGSSSHDILCTSCGTFPLSTRVPGAEECEERAVIDFVAFQDISIKRLQLQALEAPE 240
Db 181 lnpvgssshdiltctscgtfplstrvpgaeeceeravidfvafqdisikrlqlqaleape 240

QY 241 GWGPTPRAGRAALQLKLRRLTELICAGDQGALLVRLQLALRVARMPLGLSVRERFLPVH 300
Db 241 gwgptpragraalqlkrrrtelligagdgallvrlqlalrvarmplglsvrerflpvh 300

RESULT 4
ID Y42182 standard; Protein; 300 AA.
XX Y42182;
XX 17-DEC-1999 (first entry)
XX Human FLINT #1 protein sequence.
XX Human: FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
KW reperfusion-associated injury; aplastic anaemia; differentiation;
KW growth; myelodysplastic syndrome; pancytopenic condition;
KW myocardial ischaemia.
XX Homo sapiens.
XX OS
XX PN
XX PN
XX PD
XX PF
XX PF 30-MAR-1999; 99WO-US06797.
XX PR 30-MAR-1998; 98US-0079856.
XX PR 20-MAY-1998; 98US-0086074.
XX PR 09-SEP-1998; 98US-0099643.
XX PR 17-DEC-1998; 98US-0112577.
XX PR 18-DEC-1998; 98US-0112703.
XX PR 18-DEC-1998; 98US-0112933.
XX PR 22-DEC-1998; 98US-0113407.
XX (ELIL) LILLY & CO ELL.
XX Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
PI Hui KY, Khartonenkov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
PI Song HY, Wang J, Wu X, Zuckerman SH;
XX WPI: 1999-591319/50.
XX N-PSDB; Z25375.
XX Use of mature FLINT for treating acute liver failure, inflammation,
PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
PT and proinflammatory activity
XX Claim 30; Fig 1; 99pp; English.
XX The present invention describes therapeutic applications of mature FLINT
CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
CC (mFLINT), which is a member of the tumour necrosis factor receptor
CC superfamily, is used for treating acute liver failure, inflammation of
CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
CC injury or disorder such as hypercoagulation (including use with
CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
CC or disorder, Type I diabetes, cancer, cell damage or damage to an
CC innocent bystander tissue that is induced by a chemotherapeutic agent or
CC therapeutic irradiation, treating haematopoietic progenitor cells that
CC have been exposed to therapeutic radiation or chemotherapy, aplastic
CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
CC also used for promoting the growth or differentiation of a haematopoietic
CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
CC resulting from abnormal myocardial ischaemia. The present sequence
CC represents human FLINT.
XX Sequence 300 AA;
SQ

Query Match 100.0%; Score 300; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 5.6e-270; Indels 0; Gaps 0;
Matches 300; Conservative 0; Mismatches 0;
QY 1 MRALEGPGLSLCLLALPALLPVPVAVRGVAETPTYPWRDAETGERLVCAQCPCPCTFVQR 60
Db 1 mralegpglsllclvlalpallvpavrgvaetptyprdaetgerlvcaqcpptfvqr 60
QY 61 PQRDSPTTCGCPGPRHYTFWNYLERCRYCNVLCGEREEERACHATINRACRGTGFF 120
Db 61 pqrdspttcgpcprhytqfwnylercrycnvlgereeerachathnracrctgff 120
QY 121 AHAGFCLRHASCPGAGVIAPCTPSONTCQPCPCPTFFSASSSSBOCPHNRNCTALGLA 180
Db 121 ahagfclrhscppgagviapctpsntqtcpcppptfsassssseqcpnrcntalglia 180
QY 181 LNVPGSSSHDTLCTCTGFPFLSTRVPGAEECERAVIDFVAFQDISIKRLQRLQALEAPE 240
Db 181 lnvpgssshdtlctctgfpflstrvpgaeeeceravidfvafqdisikrllqalleape 240
QY 241 GWGPTPRAGRAALQLKLRRLTELICAGDQGALLVRLQLALRVARMPLGLSVRERFLPVH 300
Db 241 gwgptpragraalqlkrrrtelligagdgallvrlqlalrvarmplglsvrerflpvh 300

RESULT 5
YI7479
ID YI7479 standard; Protein; 300 AA.
XX AC
XX YI7479;
XX DT 02-AUG-1999 (first entry)
XX DE Mammalian tumour necrosis factor receptor OPG-2.
XX KW Tumour necrosis factor receptor; TNF receptor; OPG-2; Paget's disease;
KW osteopenic disorder; osteoclast activity; primary osteoporosis;
KW hyperglycaemia; osteolytic metastasis; immune response; cancer.
XX OS Mammalia.
XX PN WO9926977-A1.
XX PD 03-JUN-1999.
XX PF 24-NOV-1998; 98WO-US25065.
XX PR 17-FEB-1998; 98US-0074896.
XX PR 24-NOV-1997; 97US-0066446.
XX PA (BIOJ) BIOGEN INC.
XX PI Tschopp J;
XX WPI: 1999-347693/29.
XX N-PSDB; X76052.
XX New tumour necrosis factor family receptor OPG-2
XX Claim 1; Page 18; 22pp; English.

The present sequence represents a mammalian tumour necrosis factor receptor, designated OPG-2. OPG-2, is a member of the tumour necrosis factor receptor family, and can be used: (i) to raise specific antibodies (Ab), (ii) to treat osteopenic disorders associated with excessive osteoclast activity, e.g. primary osteoporosis, Paget's disease, hyperglycaemia of malignancy, or osteolytic metastases; (iii) for affinity purification of cognate ligands, and (iv) to screen for ligands (antagonists or agonists). Ab, or other OPG-2 blocking agents such as soluble forms of the protein, are used to prevent, or reduce severity of, an immune response, and for treating cancer. They can also

CC be used in diagnostic assays. The nucleic acid sequence encoding OPG-2
CC can be used as a probe to isolate related sequences from other species.
XX
SQ Sequence 300 AA;

Query Match 100.0%; Score 300; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 5.6e-270;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRALEGGLSLCLVLALPALLPVPVGVGAETPTYPWRDAETGERLVCAQCPGTFVQR 60
|||||
Db 1 mralepgglsllclvlalpallpvpvgvaetptywrdaetgerlvcaqcpptfvqr 60
|||||

Qy 61 PCRRDSPTCGCPPRHYTQFWNYLRCRYCNVLCGEREEARACHATHNRACRCRTGFF 120
|||||
Db 61 pcrrdspttcgpcprhytqfwnylercrycnvlgereearachathnracrcrtgff 120
|||||

Qy 121 AHAGFCLEHASCPPGAGVIAPGTPSQNTQCPGPGTFSASSSSSEQCPHRNCTALGLA 180
|||||
Db 121 ahagfclehascpaggviapgtpsqntqcpptfsassssseqcphrncalgl 180
|||||

Qy 181 LNVPSSSHDTLCTSGTGFPLSTRVPGAECERAVIDFVAFQDISIKRLQLQALEAPE 240
|||||
Db 181 lnvpssshdtlctsgtfplstrvpgaeeceravidfvafqdisikrlqlqaleape 240
|||||

Qy 241 GWGPTPRAGRAALQLKRRRLTELLGAQDGLLVRLLOALRVAMPGLERSVRERFLPVH 300
|||||
Db 241 gwgptpragraaalqlkrrrltellgaqdgallvrlqlarvampglersvrerflpvh 300
|||||

RESULT 6
Y06817
ID Y06817 standard; Protein; 300 AA.
XX
AC Y06817;
XX
DT 24-JUN-1999 (first entry)
XX
DE Human Dcr3 polypeptide.
XX
KW Dcr3 polypeptide; tumour necrosis factor receptor; TNFR; Fas ligand;
KW apoptosis; T cell mediated immune response; allergy; asthma; cancer;
KW rheumatoid arthritis; Crohn's disease; guest vs. host disease; human;
KW gene therapy.
XX
OS Homo sapiens.
XX
PN W09914330-A1.
XX
PD 25-MAR-1999.
XX
PF 18-SEP-1998; 98WO-US19661.
XX
PR 30-JUL-1998; 98US-0094640.
PR 18-SEP-1997; 97US-0059288.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Botstein D, Dodge KH, Goddard A, Gurney AL;
PI Kim KJ, Lawrence DA, Pitti R, Roy MA, Tumas DB;
PI Wood WI;
XX
DR WPI; 1999-244032/20.
DR N-PSDB; X32744.
XX
PT Dcr3 polypeptide related to tumor necrosis factor receptor
XX
PS Claim 5; Fig 1; 88pp; English.
XX
CC This represents a human Dcr3 polypeptide, a homologue of tumour necrosis
CC factor receptor (TNFR) polypeptide. Host cells containing a vector
CC comprising the Dcr3 nucleic acid can be used for the recombinant

CC expression of the protein. Dcr3 binds to Fas ligand, so it (or its
CC chimeras) are useful for modulating apoptosis in mammalian cells, also
CC other Fas-ligand induced activities, particularly to inhibit T cell
CC mediated immune responses, e.g. in treatment of allergy, asthma,
CC rheumatoid arthritis, Crohn's disease, guest vs. host disease etc. Dcr3
CC may also be used to identify specific binding proteins, potential
CC inhibitors. Antibodies against Dcr3 are used to treat cancer,
CC specifically of the lung and colon, also in diagnosis and for affinity
CC purification of the protein. Detecting mutations in the gene for Dcr3 is
CC also used to diagnose cancer, or predisposition to it. Dcr3 nucleic acid
CC is useful as hybridization probe to detect genomic or related sequences;
CC for chromosome and gene mapping; as source of antisense sequences; for
CC expression of recombinant Dcr3 and to generate transgenic animals (for
CC development and screening of therapeutic agents), also for in vivo or
CC ex vivo gene therapy.
XX
SQ Sequence 300 AA;

Query Match 100.0%; Score 300; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 5.6e-270;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRALEGGLSLCLVLALPALLPVPVGVGAETPTYPWRDAETGERLVCAQCPGTFVQR 60
|||||
Db 1 mralepgglsllclvlalpallpvpvgvaetptywrdaetgerlvcaqcpptfvqr 60
|||||

Qy 61 PCRRDSPTCGCPPRHYTQFWNYLRCRYCNVLCGEREEARACHATHNRACRCRTGFF 120
|||||
Db 61 pcrrdspttcgpcprhytqfwnylercrycnvlgereearachathnracrcrtgff 120
|||||

Qy 121 AHAGFCLEHASCPPGAGVIAPGTPSQNTQCPGPGTFSASSSSSEQCPHRNCTALGLA 180
|||||
Db 121 ahagfclehascpaggviapgtpsqntqcpptfsassssseqcphrncalgl 180
|||||

Qy 181 LNVPSSSHDTLCTSGTGFPLSTRVPGAECERAVIDFVAFQDISIKRLQLQALEAPE 240
|||||
Db 181 lnvpssshdtlctsgtfplstrvpgaeeceravidfvafqdisikrlqlqaleape 240
|||||

Qy 241 GWGPTPRAGRAALQLKRRRLTELLGAQDGLLVRLLOALRVAMPGLERSVRERFLPVH 300
|||||
Db 241 gwgptpragraaalqlkrrrltellgaqdgallvrlqlarvampglersvrerflpvh 300
|||||

RESULT 7
W97749
ID W97749 standard; Protein; 300 AA.
XX
AC W97749;
XX
DT 21-MAY-1999 (first entry)
XX
DE Human tumour necrosis factor receptor ZTNFR-5.
XX
KW ZTNFR-5; tumour necrosis factor receptor; TNFR; human;
KW cell maturation; bone cell regulation.
XX
OS Homo sapiens.
XX
PI Key Location/Qualifiers
PI Peptide 1..23
FT /note= "signal peptide"
FT Protein 24..300
FT /note= "mature protein"
FT Domain 24..194
FT /note= "extracellular domain"
FT Region 49..71
FT /note= "cysteine-rich pseudo-repeat 1"
FT Region 72..113
FT /note= "cysteine-rich pseudo-repeat 1"
FT Region 114..151
FT /note= "cysteine-rich pseudo-repeat 1"
FT Region 152..194

Y42184
ID Y42184 standard; Protein; 271 AA.
AC Y42184;
XX
DT 17-DEC-1999 (first entry)
XX
DE Human mFLINT #1 protein sequence.
XX
KW Human: FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
KW reperfusion-associated injury; aplastic anaemia; differentiation;
KW growth; myelodysplastic syndrome; pancytopenic condition;
KW myocardial ischaemia.
XX
OS Homo sapiens.
XX
PN W09950413-A2.
XX
PD 07-OCT-1999.
XX
PF 30-MAR-1999; 99WO-US06797.
XX
PR 30-MAR-1998; 98US-0079856.
PR 20-MAY-1998; 98US-0086074.
PR 09-SEP-1998; 98US-0099643.
PR 17-DEC-1998; 98US-0112577.
PR 18-DEC-1998; 98US-0112703.
PR 18-DEC-1998; 98US-0112933.
PR 22-DEC-1998; 98US-0113407.
XX
PA (ELIL) LILLY & CO ELI.
XX
PI Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
PI Song HV, Wang J, Wu X, Zuckerman SH;
XX
DR WPI: 1999-591319/50.
DR N-PSDB: Z25377.
XX
PT Use of mature FLINT for treating acute liver failure, inflammation,
PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
PT and proinflammatory activity
XX
PS Claim 31; Fig 3; 99pp; English.
XX
XX The present invention describes therapeutic applications of mature FLINT
CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
CC (mFLINT), which is a member of the tumour necrosis factor receptor
CC superfamily, is used for treating acute liver failure, inflammation of
CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
CC injury or disorder such as hypercoagulation (including use with
CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
CC or disorder, Type I diabetes, cancer, cell damage or damage to an
CC innocent bystander tissue that is induced by a chemotherapeutic agent or
CC therapeutic irradiation, treating haematopoietic progenitor cells that
CC have been exposed to therapeutic radiation or chemotherapy, aplastic
CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
CC also used for promoting the growth or differentiation of a haematopoietic
CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
CC resulting from abnormal myocardial ischaemia. The present sequence
CC represents human mFLINT.
XX
SQ Sequence 271 AA;

Query Match 90.3%; Score 271; DB 20; Length 271;
Best Local Similarity 100.0%; Pred. No. 3.7e-243;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

.QY 30 VAETPTYPWRDAETGERLVCACQPPGTFVQPCRRDSPTTCGCPPRHYTQFWNYLRCR 89

Db 1 vaetptypwrdaetgerlvcacqppgtfvqpcrrdspttcgcpprhytqfwnylrcr 60
QY 90 YCNVLGGEREEARACHATHNRACRRTGFFAHAGFCLEHASCPCGAGVIAPGTFSQNTQ 149
Db 61 ycnvlcgereearachathnracrcrtgffahagfclehascpbgagviapgtpsqntq 120
QY 150 COPCPPGTFSASSSSSEOCQPHRNCTALGLALNVPGSSSHDFTCTCTGCFPLSTRVPGA 209
Db 121 qcpcppgtfssssseqcqhnrctalglnvpgssshdctcctgfpflstrvpgae 180
QY 210 ECERAVIDFVAFQDISIKRLQALQALEPQGWGTPPRAGRAALQLKRRRLTELLGAQD 269
Db 181 eceravidfvafqdisikrlqlqaleapegwgtpragtaalqlkrrrlteillgaqd 240
QY 270 GALLVRLLOALRVARMPGLERSVRERFLPVH 300
Db 241 gallvrlqlarvarmpglersvrerflpvh 271
RESULT 10
Y28449
ID Y28449 standard; Protein; 245 AA.
XX
AC Y28449;
XX
DT 29-SEP-1999 (first entry)
XX
DE A human tumour necrosis factor-R2-like proteins (TR2P)-1.
XX
KW Human tumour necrosis factor-R2-like protein; TR2P; achondroplasia;
KW osteoporosis; developmental disorder; Cushing's syndrome;
KW muscular dystrophy; epilepsy; hereditary neuropathy;
KW Charcot-Marie-Tooth disease; neurofibromatosis; hypothyroidism;
KW hydrocephalus; seizure disorder; cerebral palsy; spinal bifida;
KW congenital glaucoma; cataract; sensorineural hearing loss;
KW reproductive disorder; infertility; ovulatory defect; endometriosis;
KW autoimmune disorder; ectopic pregnancy; teratogenesis; spermatogenesis;
KW immunological disorder; AIDS; Addison's disease; allergy; bronchitis;
KW atherosclerosis; diabetes mellitus; Chron's disease; lupus;
KW irritable bowel syndrome; multiple sclerosis; infection;
KW neoplastic disorder; adenocarcinoma; leukaemia; lymphoma; melanoma;
KW myeloma; sarcoma.
XX
OS Homo sapiens.
XX
PN W09931128-A2.
XX
PD 24-JUN-1999.
XX
PF 02-DEC-1998; 98WO-US25649.
XX
PR 16-DEC-1997; 97US-0991945.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Au-Young J, Bandman O, Hillman JL, Kaser MR, Tang YT;
XX
DR WPI: 1999-457916/38.
DR N-PSDB: X89503.
XX
PT New tumour necrosis factor-R2-like protein - useful in the treatment
PT of osteogenesis, developmental, reproductive, immunological and
PT neoplastic disorders
XX
PS Claim 1; Fig 1A-C; 81pp; English.
XX
CC The present sequence represents a human tumour necrosis factor-R2-like
CC protein (TR2P). The protein is used to treat and prevent osteogenesis,
CC developmental, reproductive, immunological and neoplastic disorders, and
CC also to diagnose disorders associated with TR2 protein expression. Such
CC disorders include osteogenesis disorders such as achondroplasia and
CC osteoporosis, developmental disorders such as Cushing's syndrome,

CC muscular dystrophy, and epilepsy, hereditary neuropathies such as
 CC Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism,
 CC hydrocephalus, seizure disorders such as cerebral palsy and spinal
 CC bifida, congenital glaucoma, cataract, or sensorineural hearing loss,
 CC reproductive disorders such as infertility, ovulatory defects and
 CC endometriosis, autoimmune disorders, ectopic pregnancy and teratogenesis,
 CC disruption of spermatogenesis, immunological disorders such as AIDS,
 CC Addison's disease, allergies, bronchitis, atherosclerosis, diabetes
 CC mellitus, Chron's disease, lupus and irritable bowel syndrome, multiple
 CC sclerosis, viral, fungal, helminthic, parasitic and protozoal infections,
 CC and neoplastic disorders including adenocarcinoma, leukaemia, lymphoma,
 CC melanoma, myeloma, sarcoma, and teratocarcinoma.
 XX
 SQ Sequence 245 AA;

Query Match 80.0%; Score 240; DB 20; Length 245;
 Best Local Similarity 100.0%; Pred. No. 1.7e-214;
 Matches 240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRALEGGSLCLLVLPALLPVPVAVRGVAETPTVWRDAETGERLVCAQCQPPGTFVOR 60
 DB 1 mraleppgslcllvlpallpvpavrgvaetpvtvwrdaetgerlvcaqcpptgfvqr 60
 QY 61 PCRRDPTTCGCPPRHYTFWNYLRCRYCNVLCGEREEARACHATHNRACRRTGFF 120
 DB 61 perrdpttcgpcprhytqfwnylercrycnvlgereearachathnracrtgff 120
 QY 121 AHAGFCLHASCPPGAGVIAPTGTPSNTQCQPCPGTFSASSSSSCQCPHRNCTALGLA 180
 DB 121 ahagfclhascpaggviaptgtpsqntqcpcpgtfsassssscqcpshrntalgla 180
 QY 181 LNVPGSSSHDTLCTCTGFFPLSTRVPGAECERAVIDFVAFQDISIKRLQLQALEAPE 240
 DB 181 lnpvpgssshdtlctctgffplstrvpgaeceravidfvafqdisikrlqlqaleape 240

RESULT 11
 W93585
 ID W93585 standard; Protein; 215 AA.
 AC W93585;
 XX
 XX
 DT 18-JUN-1999 (first entry)
 DE Human hAPO6 protein.
 XX
 KW Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
 KW developmental abnormality; gestational abnormality; prostate cancer;
 KW APO6; APO8; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
 KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
 KW apoptosis; human.
 OS Homo sapiens.
 XX
 XX WO9911791-A2.
 XX
 PD 11-MAR-1999.
 XX
 PF 04-SEP-1998; 98WO-US18393.
 XX
 XX 05-SEP-1997; 97US-0924634.
 XX
 XX (UNIW) UNIV WASHINGTON.
 XX
 XX Chaudhary PM;
 XX
 DR WPI; 1998-205191/17.
 DR N-PSDB; X23419.
 XX
 XX New Tumor Necrosis Factor family receptor polypeptides and ligands -
 PT useful for diagnosis and treatment of prostate cancer and
 PT developmental or gestational abnormalities

XX
 PS
 XX Claim 29; Fig 9; 156pp; English.
 CC This invention describes isolated Tumor Necrosis Factor (TNF) family
 CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
 CC their active fragments. APO4 is useful for diagnosing prostate cancer
 CC by determining levels of APO4 in an individual. Prostate cancer can also
 CC be treated using APO4 selective binding agents linked to a therapeutic
 CC moiety. APO4 polypeptides are also useful for identifying selective
 CC binding agents, useful in diagnosis/treatment of disease by binding of
 CC agents to the polypeptide/active fragment which is extracellular, or
 CC expressed on the cell surface. The binding is preferably performed in
 CC vivo. APO4 polypeptides/ active fragments are also useful for screening
 CC for agonists and antagonists by binding and observing the change in APO4
 CC activity. Effective pharmacological agents useful in diagnosis or
 CC treatment of disease are also identified using APO4 polypeptides or
 CC fragments and APO4 signal transducer molecules that specifically interact
 CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4
 CC activity. The method is performed in vivo or in vitro. APO polypeptides
 CC are all useful as immunogens for preparing antibodies. APO4 is also
 CC useful for diagnosis/treatment of developmental or gestational
 CC abnormalities. APO8 was transfected to human breast carcinoma cell line
 CC MCF-7, and induced apoptosis.

XX Sequence 215 AA;

Query Match 71.7%; Score 215; DB 20; Length 215;
 Best Local Similarity 100.0%; Pred. No. 2.1e-191;
 Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 86 ERCRYCNVLCGEREEARACHATHNRACRRTGFFAHAGFCLHASCPPGAGVIAPTGFS 145
 DB 1 ercrycnvlgereearachathnracrtgffahagfclhascpaggviaptgfs 60
 QY 146 QNTQCQPCPGTFSASSSSSCQCPHRNCTALGLALNVPGSSSHDTLCTCTGFFPLSTRV 205
 DB 61 qntqcqpcpgtfsassssscqcpshrntalgla lnpvpgssshdtlctctgffplstrv 120
 QY 206 PGAEECERAVIDFVAFQDISIKRLQLQALEAPEGCGTPPRAGRAALQKLRRRLTELL 265
 DB 121 pgaeceravidfvafqdisikrlqlqaleapegwgtppragraalqlklrrrltell 180
 QY 266 GAQDGALLVRLLOALRVARMPGLRSVRERFLPVH 300
 DB 181 gaqdgallvrlqlarvmpgglersvrerflpvh 215

RESULT 12
 Y77458
 ID Y77458 standard; Protein; 300 AA.
 XX
 XX Y77458;
 XX
 DT 05-JUN-2000 (first entry)
 DE Human TNF receptor-like protein, HDTEA84.
 XX
 XX TNF receptor family; tumour necrosis factor; HDTEA84; HSLJD37R;
 KW Rank-like protein; RANKL; immune disorder; inflammation; allergy;
 KW immunosuppressant; antiarthritic; antirheumatoid; antiinflammatory;
 KW dermatological; antithyroid.
 XX
 XX Homo sapiens.
 XX
 XX WO200001817-A2.
 XX
 PD 13-JAN-2000.
 XX
 PF 06-JUL-1999; 99WO-US12366.
 XX
 PR 06-JUL-1998; 98US-0110938.

QY 85 LERCRCYNVLCGERDEARACHATHNRACRRTGFFAHAGFCLHASCPCPGAGVIAPGTP 144
|||||
Db 1 lercrcynvlgereearachathnracrtgffahagfclhascpcpgagviapgt 60
QY 145 SONTQCPCPPGTFSSASSSSQCPHRCNTALGALNVPGSSSHDTLCTCTGFPPLSTR 204
|||||
Db 61 sgntqpcppgtfssassssseqcpnrctalgainvpgssshdtlctctgfpplstr 120
QY 205 VPGAECERAVIDFVAFQDISIKRLORLQALE 237
|||||
Db 121 vpgaeceravidfvafqdisikriqligale 153
RESULT 14
W63623
ID W63623 standard; Protein: 170 AA.
XX
AC W63623;
XX
DT 26-OCT-1998 (first entry)
XX
DE Human tumour necrosis factor receptor-6 beta protein.
XX
KW Human tumour necrosis factor receptor-6 alpha; TNFR-6 beta;
KW endothelial cells; keratinocytes; normal prostate; apoptosis;
KW prostate tumour tissue.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..30
FT Protein 31..170
FT note= "TNFR-6 beta"
FT Region 31..166
FT /note= "Soluble extracellular domain"
XX
XX W09830694-A2.
XX
PD 16-JUL-1998.
XX
PF 13-JAN-1998; 98WO-US00153.
XX
PR 14-JAN-1997; 97US-0035496.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ebner R, Feng P, Gentz RL, Ni J, Ruben SM, Yu G;
XX
XX WPI; 1998-399142/34.
XX
DR N-PSDB; V39086.
XX
PT Human tumour necrosis factor receptors 6-alpha and 6-beta - used in
PT the diagnosis of immune system-related disorder(s)
XX
PS Claim 20; Fig 2A-2B; 91pp; English.
XX
XX The present sequence represents the human tumour necrosis factor
CC receptor-6 beta (TNFR-6 beta) protein. The invention also provides
CC for the TNFR-6 alpha protein (W63622). TNFR-6 alpha and TNFR-6 beta are
CC members of the tumour necrosis factor receptor (TNFR) family. TNFRs
CC are expressed in endothelial cells, keratinocytes, normal prostate and
CC prostate tumour tissue. For a number of disorders of these cells,
CC particularly of the immune system, substantially altered (whether
CC increased or decreased) levels of TNFR-6 alpha and/or TNFR-6 beta gene
CC expression can be detected, therefore the TNFR-6 alpha and TNFR-6 beta
CC polypeptides, nucleic acids and antibodies are claimed to be useful in
CC the diagnosis of such disorders. Mutations of the TNFR-6 alpha and
CC TNFR-6 beta genes can also be detected. The TNFR polypeptides are
CC also claimed to be useful for identifying ligands which may be useful
CC in the treatment of apoptosis related disorders.
XX
SQ Sequence 170 AA;

Query Match 47.3%; Score 142; DB 19; Length 170;
Best Local Similarity 100.0%; Pred. No. 6.9e-124; Indels 0; Gaps 0;
Matches 142; Conservative 0; Mismatches 0;
QY 1 MRALEGSLSLCLVLALPALLPVPVARGVAETPTYPWRDAETGERLVCACQCPGTFVOR 60
|||||
Db 1 mralegslslclvlalpallpvpavrgvaetptypwrdaetgerlvcacqcpgtfvqr 60
QY 61 PCRRDSPPTCGPCPPRHHTYQFHWYLERCRYCNVLCGEREEERACHATHNRACRRTGFF 120
|||||
Db 61 pcrrdspttcgpcpprrhytqfwnylercrcynvlgereeeearachathnracrtgff 120
QY 121 AHAGFCLHASCPCPGAGVIAPG 142
|||||
Db 121 ahagfclhascpcpgagviapg 142
RESULT 15
Y42185
ID Y42185 standard; Protein: 273 AA.
XX
AC Y42185;
XX
DT 17-DEC-1999 (first entry)
XX
DE Human mFLINT #2 protein sequence.
XX
KW Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
KW reperfusion-associated injury; aplastic anaemia; differentiation;
KW growth; myelodysplastic syndrome; pancytopenic condition;
KW myocardial ischaemia.
XX
OS Homo sapiens.
XX
PN W09950413-A2.
XX
PD 07-OCT-1999.
XX
PF 30-MAR-1999; 99WO-US06797;
XX
PR 30-MAR-1998; 98US-0079856.
PR 20-MAY-1998; 98US-0086074.
PR 09-SEP-1998; 98US-0099643.
PR 17-DEC-1998; 98US-0112577.
PR 18-DEC-1998; 98US-0112703.
PR 18-DEC-1998; 98US-0112933.
PR 22-DEC-1998; 98US-0113407.
XX
PA (ELIL) LILLY & CO ELI.
XX
XX Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
PI Hui KY, Kharitonkov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
PI Song HY, Wang J, Wu X, Zuckerman SH;
XX
XX WPI; 1999-591319/50.
DR N-PSDB; Z25378.
XX
XX Use of mature FLINT for treating acute liver failure, inflammation,
PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
PT and proinflammatory activity
XX
PS Example 2; Fig 4; 99pp; English.
XX
XX The present invention describes therapeutic applications of mature FLINT
CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
CC (mFLINT), which is a member of the tumour necrosis factor receptor
CC superfamily, is used for treating acute liver failure, inflammation of
CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated

CC injury or disorder such as hypercoagulation (including use with
 CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
 CC or disorder, Type I diabetes, cancer, cell damage or damage to an
 CC innocent bystander tissue that is induced by a chemotherapeutic agent or
 CC therapeutic irradiation, treating haematopoietic progenitor cells that
 CC anaemia, myelodysplastic syndrome or chemotherapy, aplastic
 CC also used for promoting the growth or a pancytopenic condition. mFLINT is
 CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
 CC resulting from abnormal myocardial ischaemia. The present sequence
 CC represents human mFLINT.
 XX
 SQ Sequence 273 AA;

Query Match 43.3%; Score 130; DB 20; Length 273;
 Best Local Similarity 100.0%; Pred. No. 1.3e-112;
 Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 113 CRCRTGFFAHAGFCLEHASCPGAGVIAPGTPSQNTQCPCPPGTFSSSSSECCQPHR 172
 DB 86 CRCRTGFFAHAGFCLEHASCPGAGVIAPGTPSQNTQCPCPPGTFSSSSSECCQPHR 145
 QY 173 NCTALGLALNVPGSSSHDTLCTCTGTFPLSTRVPGAECERAVIDFVAFQDISIKRLQL 232
 DB 146 nctalglalnvpvgssshdtlctctgfpflstrvpgaeceravidfvafqdisikrlql 205
 QY 233 LQALEAPEGW 242
 DB 206 lqaleapegw 215

RESULT 16
 Y42183
 ID Y42183 standard; Protein; 302 AA.

XX AC Y42183;
 XX DT 17-DEC-1999 (first entry)
 XX DE Human FLINT #2 protein sequence.
 XX Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
 KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
 KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
 KW reperfusion-associated injury; aplastic anaemia; differentiation;
 KW growth; myelodysplastic syndrome; pancytopenic condition;
 KW myocardial ischaemia.

XX OS Homo sapiens.
 XX PN W09950413-A2.
 XX PD 07-OCT-1999.
 XX PF 30-MAR-1999; 99WO-US06797.

XX PR 30-MAR-1998; 98US-0079856.
 XX PR 20-MAY-1998; 98US-0086074.
 XX PR 09-SEP-1998; 98US-0099643.
 XX PR 17-DEC-1998; 98US-0112577.
 XX PR 18-DEC-1998; 98US-0112703.
 XX PR 18-DEC-1998; 98US-0112933.
 XX PR 22-DEC-1998; 98US-0113407.

XX (ELIL) LILLY & CO ELI.

XX Buml TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
 PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
 PI Song HY, Wang J, Wu X, Zuckerman SH;
 XX WPI; 1999-591319/50.
 DR N-PSDB; 225376.

XX Use of mature FLINT for treating acute liver failure, inflammation,
 PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
 PT and proinflammatory activity
 PS Example 2; Fig 2; 99pp; English.

XX The present invention describes therapeutic applications of mature FLINT

CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
 CC (mFLINT), which is a member of the tumour necrosis factor receptor
 CC superfamily, is used for treating acute liver failure, inflammation of
 CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
 CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
 CC injury or disorder such as hypercoagulation (including use with
 CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
 CC or disorder, Type I diabetes, cancer, cell damage or damage to an
 CC innocent bystander tissue that is induced by a chemotherapeutic agent or
 CC therapeutic irradiation, treating haematopoietic progenitor cells that
 CC have been exposed to therapeutic radiation or chemotherapy, aplastic
 CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
 CC also used for promoting the growth or a pancytopenic condition. mFLINT is
 CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
 CC resulting from abnormal myocardial ischaemia. The present sequence
 CC represents human FLINT.
 XX

SQ Sequence 302 AA;

Query Match 43.3%; Score 130; DB 20; Length 302;
 Best Local Similarity 100.0%; Pred. No. 1.3e-112;
 Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 CRCRTGFFAHAGFCLEHASCPGAGVIAPGTPSQNTQCPCPPGTFSSSSSECCQPHR 172
 DB 115 CRCRTGFFAHAGFCLEHASCPGAGVIAPGTPSQNTQCPCPPGTFSSSSSECCQPHR 174
 QY 173 NCTALGLALNVPGSSSHDTLCTCTGTFPLSTRVPGAECERAVIDFVAFQDISIKRLQL 232
 DB 175 nctalglalnvpvgssshdtlctctgfpflstrvpgaeceravidfvafqdisikrlql 234
 QY 233 LQALEAPEGW 242
 DB 235 lqaleapegw 244

RESULT 17
 Y41111
 ID Y41111 standard; Protein; 408 AA.

XX AC Y41111;
 XX DT 17-JAN-2000 (first entry)
 XX DE Human TANGO 129 (T129) mature protein.

XX KW Tumor necrosis factor receptor; TNF; T129 protein; TANGO 129; human;
 KW cellular process; immunological disorder; abnormal lymphoid development;
 KW thymic development; T-cell mediated immune response; humoral B cell;
 KW skeletal muscle disorder; drug screening.

XX OS Homo sapiens.

XX PN W09952924-A1.

XX PD 21-OCT-1999.

XX PF 08-APR-1999; 99WO-US07832.

XX PR 09-APR-1998; 98US-0057951.

XX PA (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.

XX PI Holtzman D;

XX WPI; 1999-620368/53.
 DR N-PSDB; 223121.
 XX
 XX New isolated tumor necrosis factor receptor member used to develop
 PT products for treating, e.g. immunological disorders or disorders of the
 PT skeletal muscle
 XX
 XX Claim 8; Page 114-115; 118pp; English.
 PS
 XX The invention provides an isolated human tumor necrosis factor (TNF)
 CC receptor member, T129 (also referred as TANGO 129). The T129 polypeptide
 CC can be expressed by standard recombinant methodology. The T129
 CC polypeptides are useful as modulating agents in regulating a variety of
 CC cellular processes. Agents or modulators which have a stimulatory or
 CC inhibitory effect on T129 activity (e.g. T129 gene expression) as
 XX identified by a screening assay can be administered to individuals to
 CC treat (prophylactically or therapeutically) disorders, e.g. an
 CC immunological disorder associated with aberrant T129 activity, disorders
 CC associated with abnormal lymphoid and/or thymic development, T-cell
 CC mediated immune response, T-cell dependent help for B cells, and abnormal
 CC humoral B cell activity, and possibly disorders of the skeletal muscle.
 CC The products can also be used for detection, diagnosis, drug screening
 CC and production of transgenic animals. The present sequence represents
 XX the T129 mature polypeptide.
 SQ Sequence 408 AA;

Query Match 3.0%; Score 9; DB 20; Length 408;
 Best Local Similarity 100.0%; Pred. No. 2.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 152 PCPPGTGFS 160
 |||||
 Db 28 pcpggtf 36

RESULT 18

Y41110
 ID Y41110 standard; Protein; 430 AA.
 XX
 AC Y41110;
 XX
 DT 17-JAN-2000 (first entry)
 XX
 XX Human TANGO 129 (T129) protein.
 DE
 XX
 KW Tumor necrosis factor receptor; TNF; T129 protein; TANGO 129; human;
 KW cellular process; immunological disorder; abnormal lymphoid development;
 KW thymic development; T-cell mediated immune response; humoral B cell;
 KW skeletal muscle disorder; drug screening.
 XX
 OS Homo sapiens.
 XX
 XX W09952924-A1.
 XX
 XX 21-OCT-1999.
 XX
 XX 08-APR-1999; 99WO-US07832.
 PF
 XX
 XX 09-APR-1998; 98US-0057951.
 PR
 XX
 XX (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.
 PA
 XX
 XX Holtzman D;
 PI
 XX
 XX WPI; 1999-620368/53.
 DR
 XX
 XX N-PSDB; 223120.
 DR
 XX
 XX New isolated tumor necrosis factor receptor member used to develop
 PT products for treating, e.g. immunological disorders or disorders of the
 PT skeletal muscle

XX Claim 8; Fig 1A-B; 118pp; English.
 PS
 XX The invention provides an isolated human tumor necrosis factor (TNF)
 CC receptor member, T129 (also referred as TANGO 129). The T129 polypeptide
 CC can be expressed by standard recombinant methodology. The T129
 CC polypeptides are useful as modulating agents in regulating a variety of
 CC cellular processes. Agents or modulators which have a stimulatory or
 CC inhibitory effect on T129 activity (e.g. T129 gene expression) as
 XX identified by a screening assay can be administered to individuals to
 CC treat (prophylactically or therapeutically) disorders, e.g. an
 CC immunological disorder associated with aberrant T129 activity, disorders
 CC associated with abnormal lymphoid and/or thymic development, T-cell
 CC mediated immune response, T-cell dependent help for B cells, and abnormal
 CC humoral B cell activity, and possibly disorders of the skeletal muscle.
 CC The products can also be used for detection, diagnosis, drug screening
 CC and production of transgenic animals. The present sequence represents
 XX the T129 polypeptide.
 SQ Sequence 430 AA;

Query Match 3.0%; Score 9; DB 20; Length 430;
 Best Local Similarity 100.0%; Pred. No. 2.2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 152 PCPPGTGFS 160
 |||||
 Db 50 pcpggtf 58

RESULT 19

Y70785
 ID Y70785 standard; Protein; 430 AA.

XX Y70785;

XX
 DT 31-JUL-2000 (first entry)

XX Human tumour necrosis factor receptor related gene 12 protein.

XX Human tumour necrosis factor receptor related gene 12; TR12; cytostatic;
 KW antiinfective; anti-human immunodeficiency virus; nephrotropic; cancer;
 KW antiarteriosclerotic; analgesic; cardiac; immunosuppressant; virucide;
 KW anti-allergic; antiinflammatory; hepatotropic; antidiabetic; vulnerary;
 KW anti-asthmatic; antipsoriatic; antibacterial; antiulcer; angiogenic;
 KW clone HMU445; microbial infection; nephritis; bone disease; pain;
 KW atherosclerosis; cardiovascular disorder; allergy; inflammation; asthma;
 KW neurodegenerative disease; Alzheimer's disease; graft rejection;
 KW liver disease; autoimmune disease; cardiomyopathy; diabetes; influenza;
 KW psoriasis; glomerulonephritis; septic shock; ulcerative colitis;
 KW angiogenesis; haematopoiesis; wound healing.

XX Homo sapiens.

Key	Location/Qualifiers
Peptide	1..25
	/label= Leader_sequence
Protein	26..430
	label= TR12
Domain	26..164
	/label= TR12_extracellular_domain
Domain	48..71
	/note= "TR12 cysteine rich domain"
Domain	165..181
	/label= TR12_transmembrane_domain
Domain	182..430
	/label= TR12_intracellular_domain

XX W0200023572-A1.

XX 27-APR-2000.

XX

PF 19-OCT-1999; 99WO-US24413.
 XX
 XX
 PR 20-OCT-1998; 98US-0104950.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 PI NI J, Ruben SM;
 XX
 XX WPI: 2000-339675/29.
 DR N-PSDB; D00061.
 XX
 XX
 DR Tumor necrosis factor receptor related gene 12 (TRL2) polynucleotides
 PT and polypeptides, useful for prevention, treatment, and diagnosis of
 PT various diseases such as cancer, atherosclerosis, pain, and
 PT cardiovascular disorders -
 XX
 XX Claim 11; Fig 1; 250pp; English.
 PS
 XX The present sequence is the TRL2 (tumour necrosis factor receptor (TNFR)
 CC related gene 12) protein. TRL2 is expressed in peripheral blood
 CC lymphocytes, spleen, colon, thymus, testis, and skeletal muscle tissues.
 CC TRL2 gene and protein are useful for treatment of diseases including,
 CC cancer, microbial infections (including human immunodeficiency virus
 CC (HIV)), nephritis, bone diseases, atherosclerosis, pain, cardiovascular
 CC disorders, acquired immune deficiency syndrome (AIDS), allergy,
 CC inflammation, neurodegenerative diseases (e.g. Alzheimer's, Parkinson's),
 CC graft rejection, liver disease, autoimmune disease, cardiomyopathy, shock
 CC diabetes, influenza, asthma, psoriasis, glomerulonephritis, septic shock
 CC and ulcerative colitis. They may also be useful for promoting
 CC angiogenesis, regulating haematopoiesis, wound healing, and has various
 CC diagnostic and research applications.
 XX
 SQ Sequence 430 AA;

Query Match 3.0%; Score 9; DB 21; Length 430;
 Best Local Similarity 100.0%; Pred. No. 2.2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 PCPPTGTFSA 160
 |||||
 DB 50 pcpptgtsa 58

RESULT 20
 Y24805
 ID Y24805 standard; Protein; 27 AA.
 XX
 AC Y24805;
 XX
 DT 24-AUG-1999 (first entry)
 XX
 DE Rat HFGAN72 receptor ligand Lig 72b.
 XX
 KW HFGAN72 receptor ligand; neurotensin receptor; depression; anxiety;
 KW obsessive compulsive disorder; affective neurosis; dysthymic disorder;
 KW behaviour disorder; mood disorder; epilepsy; psychosocial dysfunction;
 KW feeding disorder; Cushing's syndrome; basophil adenoma; prolactinoma;
 KW hypopituitarism; hypophysis adenoma; hypothalamic disease; amenorrhea;
 KW Froehlich's syndrome; Kallman's syndrome; growth hormone deficiency;
 KW dwarfism; gigantism; acromegaly; migraine; acute pain; infection;
 KW sleep disorder; insomnia; jet lag; neurodegenerative disorder.
 XX
 OS Rattus sp.
 XX
 XX WO9930670-A2.
 XX
 XX 24-JUN-1999.
 XX
 XX 15-DEC-1998; 98WO-IB02143.
 PF
 XX 16-DEC-1997; 97US-0069785.
 PR
 XX 15-DEC-1997; 97US-0069459.
 PR

(SMIK) SMITHKLINE BEECHAM PLC.
 Hagan JJ, Kennett GA, Patel SR, Piper D, Smith MI;
 Terrett JA, Upton N;
 WPI: 1999-395079/33.
 Methods to promote or inhibit HFGAN72 receptor ligand interaction
 Claim 1; Fig 4; 68pp; English.
 The present invention describes methods for the treatment of subjects
 having need to promote or inhibit the interaction of an HFGAN72 receptor
 ligand and an HFGAN72 receptor. A method for the treatment of a subject
 having need to promote the interaction of an HFGAN72 receptor ligand and
 HFGAN72 receptor comprises administering an agonist that activates the
 interaction of an HFGAN72 receptor ligand and HFGAN72 receptor. The
 diseases that can be treated using the method includes depression,
 anxiety, obsessive compulsive disorder, affective neurosis, depressive
 neurosis, dysthymic disorder, behaviour disorder, mood disorder,
 epilepsy, seizure disorder, psychosocial dysfunction, sex disorder,
 disturbed biological and circadian rhythms, feeding disorders, Cushing's
 syndrome, basophil adenoma, prolactinoma, hypopituitarism, hypophysis
 adenoma, hypothalamic diseases, Froehlich's syndrome, hypothalamic
 hypogonadism, Kallman's syndrome, amenorrhea, hypothalamic disorders of
 growth hormone deficiency, dwarfism, gigantism, acromegaly, sleep
 disturbances associated with neurological disorders, migraine, acute
 pain, burn pain, neuropathic pain, back pain, pain related to infection,
 sleep disorders, insomnia, jet lag, and other neurodegenerative
 disorders. The present sequence represents an HFGAN72 receptor ligand
 for use in the method of the invention.

XX Sequence 27 AA;

Query Match 2.7%; Score 8; DB 20; Length 27;
 Best Local Similarity 100.0%; Pred. No. 1.7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLQA 235
 |||||
 DB 9 rllqllqa 16

RESULT 21
 Y24808
 ID Y24808 standard; Protein; 28 AA.
 XX
 AC Y24808;
 XX
 DT 24-AUG-1999 (first entry)
 XX
 DE Mouse HFGAN72 receptor ligand Lig 72b.
 XX
 KW HFGAN72 receptor ligand; neurotensin receptor; depression; anxiety;
 KW obsessive compulsive disorder; affective neurosis; dysthymic disorder;
 KW behaviour disorder; mood disorder; epilepsy; psychosocial dysfunction;
 KW feeding disorder; Cushing's syndrome; basophil adenoma; prolactinoma;
 KW hypopituitarism; hypophysis adenoma; hypothalamic disease; amenorrhea;
 KW Froehlich's syndrome; Kallman's syndrome; growth hormone deficiency;
 KW dwarfism; gigantism; acromegaly; migraine; acute pain; infection;
 KW sleep disorder; insomnia; jet lag; neurodegenerative disorder.
 XX
 OS Mus sp.
 XX
 XX WO9930670-A2.
 XX
 XX 24-JUN-1999.
 XX
 XX 15-DEC-1998; 98WO-IB02143.
 PF
 XX 16-DEC-1997; 97US-0069785.
 PR
 XX 16-DEC-1997; 97US-0069785.
 PR

PR 15-DEC-1997; 97US-0069459.
XX (SMIK) SMITHKLINE BEECHAM PLC.
XX Hagan JJ, Kennett GA, Patel SR, Piper D, Smith MI;
PI Terrett JA, Upton N;
XX WPI; 1999-395079/33.
XX Methods to promote or inhibit HFGAN72 receptor ligand interaction
PT Claim 1; Fig 5; 68pp; English.
XX
XX The present invention describes methods for the treatment of subjects
CC having need to promote or inhibit the interaction of an HFGAN72 receptor
CC ligand and an HFGAN72 receptor. A method for the treatment of a subject
CC having need to promote the interaction of an HFGAN72 receptor ligand and
CC HFGAN72 receptor comprises administering an agonist that activates the
CC interaction of an HFGAN72 receptor ligand and HFGAN72 receptor. The
CC diseases that can be treated using the method includes depression,
CC anxiety, obsessive compulsive disorder, affective neurosis, depressive
CC neurosis, dysthymic disorder, behaviour disorder, mood disorder,
CC epilepsy, seizure disorder, psychosexual dysfunction, sex disorder,
CC disturbed biological and circadian rhythms, feeding disorders, Cushing's
CC syndrome, basophil adenoma, prolactinoma, hypopituitarism, hypophysis
CC adenoma, hypothalamic diseases, Froehlich's syndrome, hypothalamic
CC hypogonadism, Kallman's syndrome, amenorrhea, hypothalamic disorders of
CC growth hormone deficiency, dwarfism, gigantism, acromegaly, sleep
CC disturbances associated with neurological disorders, migraine, acute
CC pain, burn pain, neuropathic pain, back pain, pain related to infection,
CC sleep disorders, insomnia, jet lag, and other neurodegenerative
CC disorders. The present sequence represents an HFGAN72 receptor ligand
CC for use in the method of the invention.
XX
SQ Sequence 28 AA;

Query Match 2.7%; Score 8; DB 20; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLLQA 235
Db 10 rlrlllqa 17
|||||

RESULT 22
Y24801
ID Y24801 standard; Protein; 28 AA.
XX AC Y24801;
XX DT 24-AUG-1999 (first entry)
XX DE Human HFGAN72 receptor ligand Lig 72b.
XX
XX HFGAN72 receptor ligand; neuropeptide receptor; depression; anxiety;
KW obsessive compulsive disorder; affective neurosis; dysthymic disorder;
KW behaviour disorder; mood disorder; epilepsy; psychosexual dysfunction;
KW feeding disorder; Cushing's syndrome; basophil adenoma; prolactinoma;
KW hypopituitarism; hypophysis adenoma; hypothalamic disease; amenorrhea;
KW Froehlich's syndrome; Kallman's syndrome; growth hormone deficiency;
KW dwarfism; gigantism; acromegaly; migraine; acute pain; infection;
KW sleep disorder; insomnia; jet lag; neurodegenerative disorder.
XX
XX Homo sapiens.
OS
XX WO9930670-A2.
XX PN 24-JUN-1999.
XX PD 15-DEC-1998; 98WO-IB02143.
XX PF

PR 16-DEC-1997; 97US-0069785.
XX 15-DEC-1997; 97US-0069459.
XX (SMIK) SMITHKLINE BEECHAM PLC.
XX Hagan JJ, Kennett GA, Patel SR, Piper D, Smith MI;
PI Terrett JA, Upton N;
XX WPI; 1999-395079/33.
XX Methods to promote or inhibit HFGAN72 receptor ligand interaction
PT Claim 1; Fig 2; 68pp; English.
XX
XX The present invention describes methods for the treatment of subjects
CC having need to promote or inhibit the interaction of an HFGAN72 receptor
CC ligand and an HFGAN72 receptor. A method for the treatment of a subject
CC having need to promote the interaction of an HFGAN72 receptor ligand and
CC HFGAN72 receptor comprises administering an agonist that activates the
CC interaction of an HFGAN72 receptor ligand and HFGAN72 receptor. The
CC diseases that can be treated using the method includes depression,
CC anxiety, obsessive compulsive disorder, affective neurosis, depressive
CC neurosis, dysthymic disorder, behaviour disorder, mood disorder,
CC epilepsy, seizure disorder, psychosexual dysfunction, sex disorder,
CC disturbed biological and circadian rhythms, feeding disorders, Cushing's
CC syndrome, basophil adenoma, prolactinoma, hypopituitarism, hypophysis
CC adenoma, hypothalamic diseases, Froehlich's syndrome, hypothalamic
CC hypogonadism, Kallman's syndrome, amenorrhea, hypothalamic disorders of
CC growth hormone deficiency, dwarfism, gigantism, acromegaly, sleep
CC disturbances associated with neurological disorders, migraine, acute
CC pain, burn pain, neuropathic pain, back pain, pain related to infection,
CC sleep disorders, insomnia, jet lag, and other neurodegenerative
CC disorders. The present sequence represents an HFGAN72 receptor ligand
CC for use in the method of the invention.
XX
SQ Sequence 28 AA;

Query Match 2.7%; Score 8; DB 20; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLLQA 235
Db 10 rlrlllqa 17
|||||

RESULT 23
Y41113
ID Y41113 standard; peptide; 40 AA.
XX AC Y41113;
XX DT 17-JAN-2000 (first entry)
XX DE TNFR/NGFR cysteine-rich domain of T129 polypeptide.
XX
XX Tumor necrosis factor receptor; TNF; T129 protein; TANGO 129; human;
KW cellular process; immunological disorder; abnormal lymphoid development;
KW thymic development; T-cell mediated immune response; humoral B cell;
KW skeletal muscle disorder; drug screening.
XX
XX Homo sapiens.
OS
XX WO9952924-A1.
XX PN 21-OCT-1999.
XX PD 08-APR-1999; 99WO-US07832.
XX PF 09-APR-1998; 98US-0057951.
XX PR (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.
XX PA

XX Holtzman D;
 XX WPI; 1999-620368/53.
 XX New isolated tumor necrosis factor receptor member used to develop
 PT products for treating, e.g. immunological disorders or disorders of the
 PT skeletal muscle -
 XX
 XX Example 3; Fig 2; 118pp; English.
 XX
 CC The invention provides an isolated human tumor necrosis factor (TNF)
 CC receptor member, T129 (also referred as TANGO 129). The T129 polypeptide
 CC can be expressed by standard recombinant methodology. The T219
 CC polypeptides are useful as modulating agents in regulating a variety of
 CC cellular processes. Agents or modulators which have a stimulatory or
 CC inhibitory effect on T129 activity (e.g. T129 gene expression) as
 CC identified by a screening assay can be administered to individuals to
 CC treat (prophylactically or therapeutically) disorders, e.g. an
 CC immunological disorder associated with aberrant T129 activity, disorders
 CC associated with abnormal lymphoid and/or thymic development, T-cell
 CC mediated immune response, T-cell dependent help for B cells, and abnormal
 CC humoral B cell activity, and possibly disorders of the skeletal muscle.
 CC The products can also be used for detection, diagnosis, drug screening
 CC and production of transgenic animals. The present sequence represents
 CC a TNFR/NGFR cysteine-rich domain of T129 polypeptide.
 XX
 SQ Sequence 40 AA;
 Query Match 2.7%; Score 8; DB 20; Length 40;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 153 CPPGTFFSA 160
 Db 1 cpggtfss 8
 |||||
 |||||
 RESULT 24
 W61383
 ID W61383 standard; Protein; 123 AA.
 XX
 AC W61383;
 XX
 DT 02-OCT-1998 (first entry)
 XX
 DE Mouse HFGAN72 receptor protein.
 XX
 KW HFGAN72 receptor; eating disorders; renal disease; heart failure;
 KW sexual dysfunction; HIV; cancer; Parkinson's disease; urinary retention;
 KW osteoporosis; angina pectoris; myocardial infarction; psychotic disorder;
 KW neurological disorder.
 XX
 OS Mus sp.
 FH Key Location/Qualifiers
 FT Region 33..65
 FT Region /note= "Ligand 72A"
 FT Region 69..96
 FT Region /note= "Ligand 72B"
 XX
 PN EP849361-A2.
 XX
 PD 24-JUN-1998.
 XX
 PF 17-DEC-1997; 97EP-0310216.
 XX
 PR 26-SEP-1997; 97US-0939093.
 PR 17-DEC-1996; 96US-0033604.
 PR 19-MAR-1997; 97US-0820519.
 PR 02-JUL-1997; 97US-0887382.
 XX

PA (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 PI Bergsma DJ, Brooks DP, Gellai M, Wilson S, Yanagisawa M;
 XX WPI; 1998-324672/29.
 DR HFGAN72 receptor ligands - and corresponding DNA, agonists,
 XX antibodies, antagonists, etc.
 PT
 PT
 XX Claim 5; Fig 5; 35pp; English.
 PS
 XX The HFGAN72 receptor protein contains two ligands whose antagonists can be
 CC used for treating obesity, diabetes, anorexia nervosa, bulimia, cachexia,
 CC chronic renal failure, renal disease, congestive heart failure, impaired
 CC glucose tolerance and sexual dysfunction. The agonist is useful for
 CC treating anorexia nervosa, bulimia and cachexia. The HFGAN72 receptor
 CC ligand is useful for treating e.g. bacterial, fungal, protozoan and viral
 CC infections, particularly infections caused by HIV-1 or HIV-2, pain,
 CC cancers, anorexia nervosa, bulimia, cachexia, obesity, diabetes, asthma,
 CC Parkinson's disease, both acute and congestive heart failure,
 CC hypotension, hypertension, urinary retention, osteoporosis, angina
 CC pectoris, myocardial infarction, ulcers, asthma, allergies, benign
 CC prostatic hypertrophy, chronic renal failure, renal disease, impaired
 CC glucose tolerance, sexual dysfunction and psychotic and neurological
 CC disorders including anxiety, schizophrenia, manic depression, delirium,
 CC dementia, severe mental retardation and dyskinesias such as Huntington's
 CC disease or Gilles de la Tourette's syndrome.
 XX
 SQ Sequence 123 AA;
 Query Match 2.7%; Score 8; DB 19; Length 123;
 Best Local Similarity 100.0%; Pred. No. 6.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 228 RLQRLLOA 235
 Db 71 rlqrlloa 78
 |||||
 |||||
 RESULT 25
 Y24806
 ID Y24806 standard; Protein; 123 AA.
 XX
 AC Y24806;
 XX
 DT 24-AUG-1999 (first entry)
 XX
 DE Mouse HFGAN72 receptor ligand.
 XX
 KW HFGAN72 receptor ligand; neuropeptide receptor; depression; anxiety;
 KW obsessive compulsive disorder; affective neurosis; dysthymic disorder;
 KW behaviour disorder; mood disorder; epilepsy; psychosexual dysfunction;
 KW feeding disorder; Cushing's syndrome; basophil adenoma; prolactinoma;
 KW hypopituitarism; hypophysis adenoma; hypothalamic disease; amenorrhea;
 KW Froehlich's syndrome; Kallman's syndrome; growth hormone deficiency;
 KW dwarfism; gigantism; acromegaly; migraine; acute pain; infection;
 KW sleep disorder; insomnia; jet lag; neurodegenerative disorder.
 XX
 OS Mus sp.
 XX
 PN W09930670-A2.
 XX
 PD 24-JUN-1999.
 XX
 PF 15-DEC-1998; 98WO-IB02143.
 XX
 PR 16-DEC-1997; 97US-0069785.
 PR 15-DEC-1997; 97US-0069459.
 XX
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX

PI Hagan JJ, Kennett GA, Patel SR, Piper D, Smith MI;
 PI Terrett JA, Upton N;
 XX WPI; 1999-395079/33.
 XX Methods to promote or inhibit HFGAN72 receptor ligand interaction
 PT Claim 1; Fig 5; 68pp; English.
 XX
 XX The present invention describes methods for the treatment of subjects
 CC having need to promote or inhibit the interaction of an HFGAN72 receptor
 CC ligand and an HFGAN72 receptor. A method for the treatment of a subject
 CC having need to promote the interaction of an HFGAN72 receptor ligand and
 CC HFGAN72 receptor comprises administering an agonist that activates the
 CC interaction of an HFGAN72 receptor ligand and HFGAN72 receptor. The
 CC diseases that can be treated using the method includes depression,
 CC anxiety, obsessive compulsive disorder, affective neurosis, depressive
 CC neurosis, dysthymic disorder, behaviour disorder, mood disorder,
 CC epilepsy, seizure disorder, psychosexual dysfunction, sex disorder,
 CC disturbed biological and circadian rhythms, feeding disorders, hypophysis
 CC syndrome, basophil adenoma, prolactinoma, hypopituitarism, hypothalamic
 CC adenoma, hypothalamic diseases, Froehlich's syndrome, hypothalamic
 CC hypogonadism, Kallman's syndrome, amenorrhea, hypothalamic disorders of
 CC growth hormone deficiency, dwarfism, gigantism, acromegaly, sleep
 CC disturbances associated with neurological disorders, migraine, acute
 CC pain, burn pain, neuropathic pain, back pain, pain related to infection,
 CC sleep disorders, insomnia, jet lag, and other neurodegenerative
 CC disorders. The present sequence represents an HFGAN72 receptor ligand
 CC for use in the method of the invention.
 XX
 SQ Sequence 123 AA;

Query Match 2.7%; Score 8; DB 20; Length 123;
 Best Local Similarity 100.0%; Pred. No. 6.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLLOA 235
 Db 71 rlqrlloa 78
 |||||

RESULT 26
 W61382
 ID W61382 standard; Protein; 130 AA.
 XX
 AC W61382;
 XX
 DT 02-OCT-1998 (first entry)
 XX
 XX Rat HFGAN72 receptor protein.
 XX
 XX HFGAN72 receptor; eating disorders; renal disease; heart failure;
 KW sexual dysfunction; HIV; cancer; Parkinson's disease; urinary retention;
 KW osteoporosis; angina pectoris; myocardial infarction; psychotic disorder;
 KW neurological disorder.
 XX
 OS Rattus sp.
 XX
 XX Key Location/Qualifiers
 FT Region 33..65
 FT /note= "Ligand 72A"
 FT Region 69..96
 FT /note= "Ligand 72B"
 XX
 XX EP849361-A2.
 PN
 XX 24-JUN-1998.
 PD
 XX
 XX 17-DEC-1997; 97EP-0310216.
 PF
 XX 26-SEP-1997; 97US-0939093.
 PR
 PR 17-DEC-1996; 96US-0033604.

PR 19-MAR-1997; 97US-0820519.
 PR 02-JUL-1997; 97US-0887382;
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX
 PI Bergsma DJ, Brooks DP, Gellai M, Wilson S, Yanagisawa M;
 XX
 XX WPI: 1998-324672/29.
 DR N-PSDB; V28139.
 DR
 XX HFGAN72 receptor ligands - and corresponding DNA, agonists,
 PT antibodies, antagonists, etc.
 XX
 PS Claim 5; Fig 4; 35pp; English.
 XX
 XX The HFGAN72 receptor protein contains two ligands whose antagonists can be
 CC used for treating obesity, diabetes, anorexia nervosa, bulimia, cachexia,
 CC chronic renal failure, renal disease, congestive heart failure, impaired
 CC glucose tolerance and sexual dysfunction. The agonist is useful for
 CC treating anorexia nervosa, bulimia and cachexia. The HFGAN72 receptor
 CC ligand is useful for treating e.g. bacterial, fungal, protozoan and viral
 CC infections, particularly infections caused by HIV-1 or HIV-2, pain,
 CC cancer, anorexia nervosa, bulimia, cachexia, obesity, diabetes, asthma,
 CC Parkinson's disease, both acute and congestive heart failure,
 CC hypotension, hypertension, urinary retention, osteoporosis, angina
 CC pectoris, myocardial infarction, ulcers, asthma, allergies, benign
 CC prostatic hypertrophy, chronic renal failure, renal disease, impaired
 CC glucose tolerance, sexual dysfunction and psychotic and neurological
 CC disorders including anxiety, schizophrenia, manic depression, delirium,
 CC dementia, severe mental retardation and dyskinesias such as Huntington's
 CC disease or Gilles de la Tourette's syndrome.
 XX
 SQ Sequence 130 AA;

Query Match 2.7%; Score 8; DB 19; Length 130;
 Best Local Similarity 100.0%; Pred. No. 6.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLLOA 235
 Db 78 rlqrlloa 85
 |||||

RESULT 27
 W50157
 ID W50157 standard; Protein; 130 AA.
 XX
 AC W50157;
 XX
 DT 09-JUL-1998 (first entry)
 XX
 XX Rat hypocretin 35.
 DE
 XX Rat; hypocretin 35; H35; treatment; neurological disease;
 KW homeostatic dysfunction;
 KW homeostatic regulatory hormone production.
 XX
 OS Rattus rattus.
 XX
 XX WO9805352-A1.
 PN
 XX 12-FEB-1998.
 PD
 XX
 XX 01-AUG-1997; 97WO-US13657.
 PF
 XX 02-AUG-1996; 96US-0023220.
 XX
 XX (SCRI) SCRIPPS RES INST.
 XX
 XX Bloom FE, Danielson PE, De Lecea L, Foye PE, Gautvik KM;
 PI Gautvik VT, Kilduff TS, Sutcliffe JG;

XX WPI: 1998-145352/13.
 DR N-PSDB; V18865.
 XX Nucleic acid encoding hypocretin of rat and mouse - useful for
 PT diagnosis and treatment of neurological disease, homeostatic
 PT dysfunction etc., also sequence for calmodulin kinase-like protein
 XX Claim 1; Page 81; 111pp; English.
 PS
 XX The present sequence is rat hypocretin 35 (H35), which is
 CC involved in lowering body temperature and reducing food intake.
 CC Modulation of the H35 receptor can be used in the treatment of
 CC neurological disease or homeostatic dysfunction, or to control
 CC homeostatic regulatory hormone production. Hypocretin proteins can
 CC be used to raise antibodies (Ab), to identify specific agonists or
 CC antagonists, in therapy, to detect Ab and to isolate cognate
 CC receptors. Oligonucleotides based on H35 cDNA can be used to detect
 CC the hypocretin gene or its RNA transcript, and as antisense agents
 CC for inhibiting gene expression. H35 cDNA can also be used for
 CC recombinant protein production. The Ab can be used to detect or
 CC quantify hypocretin proteins and as a therapeutic inhibitor.
 XX
 SQ Sequence 130 AA;

Query Match 2.7%; Score 8; DB 19; Length 130;
 Best Local Similarity 100.0%; Pred. No. 6.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 228 RLQRLQA 235
 |||||
 Db 78 rlqrlqa 85

RESULT 28
 W50158
 ID W50158 standard; Protein; 130 AA.
 XX
 AC W50158;
 XX
 DT 09-JUL-1998 (first entry)
 XX
 DE Mouse hypocretin 35.
 XX
 KW Mouse; hypocretin 35; H35; treatment; neurological disease;
 KW homeostatic dysfunction; murine;
 KW homeostatic regulatory hormone production.
 XX
 OS Mus sp.
 XX
 PN W09805352-A1.
 XX
 PD 12-FEB-1998.
 XX
 PF 01-AUG-1997; 97WO-US13657.
 XX
 PR 02-AUG-1996; 96US-0023220.
 XX
 XX (SCRI) SCRIPPS RES INST.
 XX
 PI Bloom FE, Danielson PE, De Lecea L, Foye PE, Gautvik KM;
 PI Gautvik VT, Kilduff TS, Sutcliffe JG;
 XX
 XX WPI: 1998-145352/13.
 DR N-PSDB; V18866.
 XX
 XX Nucleic acid encoding hypocretin of rat and mouse - useful for
 PT diagnosis and treatment of neurological disease, homeostatic
 PT dysfunction etc., also sequence for calmodulin kinase-like protein
 XX
 PS Claim 1; Pages 81-82; 111pp; English.
 XX

CC The present sequence is mouse hypocretin 35 (H35), which is
 CC involved in lowering body temperature and reducing food intake.
 CC Modulation of the H35 receptor can be used in the treatment of
 CC neurological disease or homeostatic dysfunction, or to control
 CC homeostatic regulatory hormone production. Hypocretin proteins can
 CC be used to raise antibodies (Ab), to identify specific agonists or
 CC antagonists, in therapy, to detect Ab and to isolate cognate
 CC receptors. Oligonucleotides based on H35 cDNA can be used to detect
 CC the hypocretin gene or its RNA transcript, and as antisense agents
 CC for inhibiting gene expression. H35 cDNA can also be used for
 CC recombinant protein production. The Ab can be used to detect or
 CC quantify hypocretin proteins and as a therapeutic inhibitor.
 XX
 SQ Sequence 130 AA;

Query Match 2.7%; Score 8; DB 19; Length 130;
 Best Local Similarity 100.0%; Pred. No. 6.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 228 RLQRLQA 235
 |||||
 Db 78 rlqrlqa 85

RESULT 29
 Y24802
 ID Y24802 standard; Protein; 130 AA.
 XX
 AC Y24802;

XX
 DT 24-AUG-1999 (first entry)
 XX
 DE Rat HFGAN72 receptor ligand.

XX
 KW HFGAN72 receptor ligand; neuropeptide receptor; depression; anxiety;
 KW obsessive compulsive disorder; affective neurosis; dysthymic disorder;
 KW behaviour disorder; mood disorder; epilepsy; psychosexual dysfunction;
 KW feeding disorder; Cushing's syndrome; basophil adenoma; prolactinoma;
 KW hypopituitarism; hypophysis adenoma; hypothalamic disease; amenorrhea;
 KW Froehlich's syndrome; Kallman's syndrome; growth hormone deficiency;
 KW dwarfism; gigantism; acromegaly; migraine; acute pain; infection;
 KW sleep disorder; insomnia; jet lag; neurodegenerative disorder.
 XX

OS Rattus sp.
 XX
 PN W09930670-A2.
 XX

PD 24-JUN-1999.

PF 15-DEC-1998; 98WO-IB02143.

PR 16-DEC-1997; 97US-0069785.

PR 15-DEC-1997; 97US-0069459.

XX (SMIK) SMITHKLINE BEECHAM PLC.

XX Hagan JJ, Kennett GA, Patel SR, Piper D, Smith MI;
 PI Terrett JA, Upton N;

XX WPI: 1999-395079/33.

DR N-PSDB; X83592.

XX Methods to promote or inhibit HFGAN72 receptor ligand interaction

XX Claim 1; Fig 4; 68pp; English.

XX The present invention describes methods for the treatment of subjects
 CC having need to promote or inhibit the interaction of an HFGAN72 receptor
 CC ligand and an HFGAN72 receptor. A method for the treatment of a subject
 CC having need to promote the interaction of an HFGAN72 receptor ligand and
 CC HFGAN72 receptor comprises administering an agonist that activates the
 CC interaction of an HFGAN72 receptor ligand and HFGAN72 receptor. The

CC diseases that can be treated using the method includes depression,
 CC anxiety, obsessive compulsive disorder, affective neurosis, depressive
 CC neurosis, dysthymic disorder, behaviour disorder, mood disorder,
 CC epilepsy, seizure disorder, psychosexual dysfunction, sex disorder,
 CC disturbed biological and circadian rhythms, feeding disorders, Cushing's
 CC syndrome, basophil adenoma, prolactinoma, hypopituitarism, hypophysis
 CC adenoma, hypothalamic diseases, Froehlich's syndrome, hypothalamic
 CC hypogonadism, Kallman's syndrome, amenorrhea, hypothalamic disorders of
 CC growth hormone deficiency, dwarfism, gigantism, acromegaly, sleep
 CC disturbances associated with neurological disorders, migraine, acute
 CC pain, burn pain, neuropathic pain, back pain, pain related to infection,
 CC sleep disorders, insomnia, jet lag, and other neurodegenerative
 CC disorders. The present sequence represents an HFGAN72 receptor ligand
 CC for use in the method of the invention.

XX
 SQ Sequence 130 AA;

Query Match 2.7%; Score 8; DB 20; Length 130;
 Best Local Similarity 100.0%; Pred. No. 6.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLQQA 235
 |||||
 Db 78 rlgrlqlqa 85

RESULT 30
 W61381
 ID W61381 standard; Protein; 131 AA.

XX AC W61381;

XX 02-OCT-1998 (first entry)

XX Human HFGAN72 receptor protein.

XX HFGAN72 receptor; eating disorders; renal disease; heart failure;
 KW sexual dysfunction; HIV; cancer; Parkinson's disease; urinary retention;
 KW osteoporosis; angina pectoris; myocardial infarction; psychotic disorder;
 KW neurological disorder.

XX OS Homo sapiens.

XX Key Location/Qualifiers
 FH Region 34..66

FT Region /note= "Ligand 72A"

FT Region 70..97

FT /note= "Ligand 72B"

XX EP849361-A2.

XX 24-JUN-1998.

XX 17-DEC-1997; 97EP-0310216.

XX 26-SEP-1997; 97US-0939093.

XX 17-DEC-1996; 96US-0033604.

XX 19-MAR-1997; 97US-0820519.

XX 02-JUL-1997; 97US-0887382.

XX (SMIK) SMITHKLINE BEECHAM CORP.

PA (SMIK) SMITHKLINE BEECHAM PLC.

XX Bergsma DJ, Brooks DP, Gellai M, Wilson S, Yanagisawa M;

XX WPI; 1998-324672/29.

XX N-PSDB; V28138.

XX HFGAN72 receptor ligands - and corresponding DNA, agonists,
 PT antibodies, antagonists, etc.

XX Claim 2; Fig 2; 35pp; English.

XX The HFGAN72 receptor protein contains two ligands whose antagonists can
 CC be used for treating obesity, diabetes, anorexia nervosa, bulimia,
 CC cachexia, chronic renal failure, renal disease, congestive heart failure,
 CC impaired glucose tolerance and sexual dysfunction. The agonist is
 CC useful for treating anorexia nervosa, bulimia and cachexia. The HFGAN72
 CC receptor ligand is useful for treating e.g. bacterial, fungal, protozoan
 CC and viral infections, particularly infections caused by HIV-1 or HIV-2,
 CC pain, cancers, anorexia nervosa, bulimia, cachexia, obesity, diabetes,
 CC asthma, Parkinson's disease, both acute and congestive heart failure,
 CC hypotension, hypertension, urinary retention, osteoporosis, angina
 CC pectoris, myocardial infarction, ulcers, asthma, allergies, benign
 CC prostatic hypertrophy, chronic renal failure, renal disease, impaired
 CC glucose tolerance, sexual dysfunction and psychotic and neurological
 CC disorders including anxiety, schizophrenia, manic depression, delirium,
 CC dementia, severe mental retardation and dyskinesias such as Huntington's
 CC disease or Gilles de la Tourette's syndrome.

XX SQ Sequence 131 AA;

Query Match 2.7%; Score 8; DB 19; Length 131;
 Best Local Similarity 100.0%; Pred. No. 6.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLQQA 235
 |||||
 Db 79 rlgrlqlqa 86

RESULT 31

Y24799

ID Y24799 standard; Protein; 131 AA.

XX AC Y24799;

XX 24-AUG-1999 (first entry)

XX Human HFGAN72 receptor ligand.

XX HFGAN72 receptor ligand; neuroleptic receptor; depression; anxiety;
 KW obsessive compulsive disorder; affective neurosis; dysthymic disorder;
 KW behaviour disorder; mood disorder; epilepsy; psychosexual dysfunction;
 KW feeding disorder; Cushing's syndrome; basophil adenoma; prolactinoma;
 KW hypopituitarism; hypophysis adenoma; hypothalamic disease; amenorrhea;
 KW Froehlich's syndrome; Kallman's syndrome; growth hormone deficiency;
 KW dwarfism; gigantism; acromegaly; migraine; acute pain; infection;
 KW sleep disorder; insomnia; jet lag; neurodegenerative disorder.

XX OS Homo sapiens.

XX WO9930670-A2.

XX 24-JUN-1999.

XX 15-DEC-1998; 98WO-IB02143.

XX 16-DEC-1997; 97US-0069785.

XX 15-DEC-1997; 97US-0069459.

XX (SMIK) SMITHKLINE BEECHAM PLC.

XX Hagan JJ, Kennett GA, Patel SR, Piper D, Smith MI;

XX Terrett JA, Upton N;

XX WPI; 1999-395079/33.

XX N-PSDB; X83590, X83591.

XX Methods to promote or inhibit HFGAN72 receptor ligand interaction
 PT Claim 1; Fig 2; 68pp; English.

XX The present invention describes methods for the treatment of subjects

CC having need to promote or inhibit the interaction of an HFGAN72 receptor
 CC ligand and an HFGAN72 receptor. A method for the treatment of a subject
 CC having need to promote the interaction of an HFGAN72 receptor ligand and
 CC HFGAN72 receptor comprises administering an agonist that activates the
 CC interaction of an HFGAN72 receptor ligand and HFGAN72 receptor. The
 CC diseases that can be treated using the method includes depression,
 CC anxiety, obsessive compulsive disorder, affective neurosis, depressive
 CC neurosis, dysthymic disorder, behaviour disorder, mood disorder,
 CC epilepsy, seizure disorder, psychosexual dysfunction, sex disorder,
 CC disturbed biological and circadian rhythms, feeding disorders, Cushing's
 CC syndrome, basophil adenoma, prolactinoma, hypopituitarism, hypophysis
 CC adenoma, hypothalamic diseases, Froehlich's syndrome, hypothalamic
 CC hypogonadism, Kallman's syndrome, amenorrhea, hypothalamic disorders of
 CC growth hormone deficiency, dwarfism, gigantism, acromegaly, sleep
 CC disturbances associated with neurological disorders, migraine, acute
 CC pain, burn pain, neuropathic pain, back pain, pain related to infection,
 CC sleep disorders, insomnia, jet lag, and other neurodegenerative
 CC disorders. The present sequence represents an HFGAN72 receptor ligand
 CC for use in the method of the invention.

SQ Sequence 131 AA;

Query Match 2.7%; Score 8; DB 20; Length 131;
 Best Local Similarity 100.0%; Pred. No. 6.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 228 RLQRLQA 235
 |||||
 Db 79 rlrqlqla 86

RESULT 32

W22601
 ID W22601 standard; Protein: 4472 AA.

AC W22601;

DT 27-FEB-1998 (first entry)

XX Tyalactone synthase ORF1 protein.

XX Tyalactone synthase gene cluster; tyIG gene; multifunctional protein;
 KW polyketide; tyalactone synthesis; antibiotic; tylosin.

XX Streptomyces fradiae.

XX Key Location/Qualifiers

FT Misc-difference 1
 FT Domain /note= "encoded by GTG"
 FT Domain 43..447
 FT Domain /note= "ketosynthase domain, KSq(s)"
 FT Domain 586..914
 FT Domain /note= "acyltransferase domain, AT(s)"
 FT Domain 954..1038
 FT Domain /note= "acyl carrier protein domain, ACP(s)"
 FT Domain 1060..1483
 FT Domain /note= "ketosynthase domain, KS1"
 FT Domain 1606..1934
 FT Domain /note= "acyltransferase domain, AT1"
 FT Domain 2199..2378
 FT Domain /note= "ketoreductase domain, KR1"
 FT Domain 2469..2552
 FT Domain /note= "acyl carrier protein domain, ACP1"
 FT Domain 2576..2999
 FT Domain /note= "ketosynthase domain, KS2"
 FT Domain 3149..3477
 FT Domain /note= "acyltransferase domain, AT2"
 FT Domain 3502..3687
 FT Domain /note= "dehydratase domain, DH2"
 FT Domain 4016..4200
 FT Domain /note= "ketoreductase domain, KR2"
 FT Domain 4302..4385

FT /note= "acyl carrier protein domain, ACP2"
 XX
 PN EP791655-A2.
 XX
 PD 27-AUG-1997.
 XX
 PF 19-FEB-1997; 97EP-0301056.
 XX
 PR 22-FEB-1996; 96US-0012078.
 XX
 PA (ELIL) LILLY & CO ELI.
 XX
 PI Dehoff BS, Kuhstoss SA, Rosteck PR, Sutton KL;
 XX WPI: 1997-418046/39.
 DR N-PSDB: T80413.
 XX
 PT DNA encoding Streptomyces fradiae tyalactone synthase domain - for
 XX production of tylosin-related polyketide compounds
 PS Claim 9; Pages 66-80; 220pp; English.

CC W22601-W22605 represent proteins encoded by the tyalactone synthase gene
 CC cluster of the invention. The gene cluster is also referred to as the
 CC tyIG gene, and was isolated from Streptomyces fradiae. These sequences
 CC are multifunctional proteins which direct the synthesis of the polyketide
 CC tyalactone, isolated from Streptomyces fradiae. Tyalactone is the basic
 CC building block of the antibiotic tylosin. The DNA sequence can be
 CC modified so as to alter the type of carboxylic acids incorporated, the
 CC number of carboxylic acids incorporated and/or the post-condensation
 CC reactions performed, thereby resulting in novel tylosin-related
 CC polyketides.

SQ Sequence 4472 AA;

Query Match 2.7%; Score 8; DB 18; Length 4472;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 257 LRRRLTEL 264

Db 4424 lrrrltel 4431

RESULT 33

P20401
 ID P20401 standard; Protein: 10 AA.

AC P20401;

DT 30-NOV-1992 (first entry)

XX Secretin precursor formation peptide 3.

XX Strong acid; digestive canal hormone; pancreas; gastrin; pepsin;
 KW insulin.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "PhSO2 ring subst. by 1, 2 or 3 alkyl or
 FT Modified-site 4 alkoxy gps."
 FT Modified-site /note= "PhSO2 ring subst. by 1, 2 or 3 alkyl or
 FT alkoxy gps."
 XX EP47997-A.
 XX
 PD 24-MAR-1982.
 XX
 PF 11-SEP-1981; 81EP-0107186.

XX 11-SEP-1980; 80JP-0125262.
 XX (EISA) EISAI KK.
 PA Uchiyama M, Sato T, Yoshino H, Tsuchiya Y, Konishi M;
 PI Tsujii M, Hisatake Y, Koiwa A;
 XX WPI; 1982-24409E/13 (24409E).
 XX Heptacosapeptide(s) - useful for high yield conversion to high
 PT purity secretin on strong acid treatment
 XX
 PS Claim 6; Page 44; 47pp; English.
 XX
 CC The sequence in P20398 is a precursor for the production of
 CC secretin. The peptide sequences given in P20399-402 are peptides which
 CC are useful in the production of this precursor. The precursor is
 CC treated with strong acid in the preparation of secretin. Secretin is
 CC one of the digestive canal hormones and is useful in promotion of
 CC pancreatic external secretin, controlling gastrin-stimulating secretin
 CC of the stomach acid, releasing insulin, stimulating secretin of pepsin
 CC and decomposing fat. It is used as a pancreatic-function examining
 CC agent and a medicine for curing duodenal ulcers etc.
 XX
 SQ Sequence 10 AA;
 Query Match 2.3%; Score 7; DB 3; Length 10;
 Best Local Similarity 100.0%; Pred. No. 6;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 228 RLQRLQ 234
 Db 1 rlqrlq 7
 RESULT 34
 P30018
 ID P30018 standard; peptide; 10 AA.
 AC P30018;
 XX
 XX 03-SEP-1992 (first entry)
 DT Intermediate in secretin synthesis.
 DE Digestive; hormone; insulin; gastrin; pancreas; duodenal ulcer;
 KW Industrial; production.
 KW Synthetic.
 OS Key Location/Qualifiers
 FH Modified-site 1
 FT /label= X-Arg-Tos
 FT /note= "X= H or benzoyloxycarbonyl"
 FT Modified-site 4
 FT /label= Arg-Tos
 FT
 XX JP58152850-A.
 PN 10-SEP-1983.
 XX
 XX 05-MAR-1982; 82JP-0034031.
 XX
 XX 05-MAR-1982; 82JP-0034031.
 XX (EISA) EISAI KK.
 PA WPI; 1983-791974/42.
 XX
 XX Decapeptide used in synthesis of secretin - contg. arginine,
 PT leucine, glycine and valine units

XX Claim 1; Page 1; 14pp; Japanese.
 PS
 XX The C-terminal amidated peptide is an intermediate in the
 CC industrial synthesis of secretin, a digestive tract hormone.
 CC Secretin has various pharmaceutical actions such as pancreatic
 CC exocrinogenic, gastrin stimulating, gastric acid secretion
 CC inhibitory, insulin releasing, pepsin secretion promoting
 CC and adipolytic action. It is used as a reagent for study
 CC of pancreatic function or as a remedy for duodenal ulcers.
 XX
 SQ Sequence 10 AA;
 Query Match 2.3%; Score 7; DB 4; Length 10;
 Best Local Similarity 100.0%; Pred. No. 6;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 228 RLQRLQ 234
 Db 1 rlqrlq 7
 RESULT 35
 Y65459
 ID Y65459 standard; Peptide; 16 AA.
 XX
 AC Y65459;
 XX
 XX 01-FEB-2000 (first entry)
 DT Estradiol activated receptor binding peptide #5.
 DE Oestrogen receptor; estrogen; estradiol; oestrogen response element;
 KW ERE; binding; biological activity; fingerprint; molecular braille;
 KW cellular braille; modulation; tamoxifen; breast cancer; ovarian cancer;
 KW menopause; osteoporosis; selective estrogen receptor modulator;
 KW identification; characterisation; classification.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 XX WO9954728-A2.
 PN 28-OCT-1999.
 PD 26-MAR-1999; 99WO-US06664.
 PF 23-APR-1998; 98US-0082756.
 PR 09-SEP-1998; 98US-0099656.
 PR 08-JAN-1999; 99US-0115345.
 XX
 XX (NOVA-) NOVALON PHARM CORP.
 PA Paige LA, Hamilton PT, Fowlkes DM, Buehrer B, Barnett T;
 PI McDonnell DP, Christensen DJ;
 XX WPI; 2000-013281/01.
 DR
 XX Methods for identifying new receptor modulators, especially estrogen
 PT modulators to treat tamoxifen refractory breast cancer -
 XX Example 1.2; Page 152; 219pp; English.
 XX
 CC The present invention describes a method for predicting the biological
 CC activity of new receptor modulating compounds (I) using novel oligomeric
 CC peptides (biokeys) which have differential abilities to bind to 2
 CC different receptor conformations. The method is used to identify new
 CC drugs that are physiological or pharmacological agonists/antagonists and
 CC that target various receptors, which are involved in certain disease
 CC conditions. The system may be used as a primary screening tool to
 CC identify hits, to classify lead compounds from a drug screen to,
 CC characterise selective oestrogen receptor modulators (SERMs) in terms of

agonist and antagonist function and to predict possible clinical effects of SERMs such as tissue and receptor specificity. The method can also be applied to the fractionation of mixtures of SERMs to determine which components are producing agonistic and antagonistic activity. The method may be used with other receptors (e.g. progesterone, dopamine and glucocorticoid, thyroid, vitamin D, beta-adrenergic, androgen, and modulators of receptor activity. Peptides comprising a LXXLL motif may be used to modulate the oestrogen receptor in treating e.g. breast and ovarian cancer and ameliorating the effects of menopause, including osteoporosis. Y65439 to Y65652 represent oestrogen receptor, estradiol receptor and oestrogen response element binding peptides given in the exemplification of the present invention. Z35740 to Z35745 represent oligonucleotides used in the exemplification of the present invention.

XX Sequence 16 AA;

Query Match 2.3%; Score 7; DB 21; Length 16;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 272 LLVRLQL 278
|||||||
DB 8 llvrllq 14

RESULT 36

P20400
ID P20400 standard; Protein; 17 AA.

XX AC P20400;

XX DT 30-NOV-1992 (first entry)

XX DE Secretin precursor formation peptide 2.

XX KW Strong acid; digestive canal hormone; pancreas; gastrin; pepsin;
insulin.

XX OS Synthetic.

Key	Location/Qualifiers
FT Modified-site 1	/note= "But protected"
FT Modified-site 2	/note= "Tos protected"
FT Modified-site 4	/note= "PhSO2 ring substd. by 1, 2 or 3 alkyl or alkoxy gps."
FT Modified-site 5	/note= "OBut protected"
FT Modified-site 6	/note= "But protected"
FT Modified-site 8	/note= "PhSO2 ring substd. by 1, 2 or 3 alkyl or alkoxy gps."
FT Modified-site 11	/note= "PhSO2 ring substd. by 1, 2 or 3 alkyl or alkoxy gps."

XX EP47997-A.

XX PD 24-MAR-1982.

XX PF 11-SEP-1981; 81EP-0107186.

XX PR 11-SEP-1980; 80JP-0125262.

XX PA (EISA) EISAI KK.

XX Uchiyama M, Sato T, Yoshino H, Tsuchiya Y, Konishi M;
PI Tsujii M, Hisatake Y, Koiwa A;

XX WPI; 1982-24409E/13 (24409E).
XX Heptacosapeptide(s) - useful for high yield conversion to high
purify secretin on strong acid treatment
XX Claim 4; Page 44; 47pp; English.

The sequence in P20398 is a precursor for the production of secretin. The peptide sequences given in P20399-402 are peptides which are useful in the production of this precursor. The precursor is treated with strong acid in the preparation of secretin. Secretin is one of the digestive canal hormones and is useful in promotion of pancreatic external secretin, controlling gastrin-stimulating secretin of the stomach acid, releasing insulin, stimulating secretin of pepsin and decomposing fat. It is used as a pancreatic-function examining agent and a medicine for curing duodenal ulcers etc.

SQ Sequence 17 AA;

Query Match 2.3%; Score 7; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 9.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLQL 234
|||||||
DB 8 rlqrlql 14

RESULT 37

P30019
ID P30019 standard; peptide; 17 AA.

XX AC P30019;

XX DT 03-SEP-1992 (first entry)

XX DE Intermediate of secretin synthesis.

XX KW Digestive; hormone; pancreatic function; duodenal ulcer.

XX OS Synthetic.

Key	Location/Qualifiers
FT Modified-site 1	/label= X-Ser-t-Bu
FT Modified-site 2	/note= "X= H or benzyloxycarbonyl"
FT Modified-site 4	/label= Arg-Tos
FT Modified-site 5	/label= Arg-Tos
FT Modified-site 6	/label= Asp-O-t-Bu
FT Modified-site 8	/label= Ser-t-Bu
FT Modified-site 11	/label= Arg-Tos
FT Modified-site 11	/label= Arg-Tos

XX JP58152849-A.

XX PD 10-SEP-1983.

XX PF 05-MAR-1982; 82JP-0034030.

XX PR 05-MAR-1982; 82JP-0034030.

XX PA (EISA) EISAI KK.

XX WPI; 1983-791973/42.

PT Hepta:deca:peptide(s), useful intermediates to secretin - a
 PT digestive tract hormone useful for treating duodenal ulcers
 XX
 PS Claim 1; Page 1; 13pp; Japanese.

XX The C-terminal amidated peptide is an intermediate in the
 CC industrial synthesis of secretin, a digestive tract hormone.
 CC Secretin has various pharmaceutical actions such as pancreatic
 CC exocrinogenic, gastrin stimulating, gastric acid secretion
 CC inhibitory, insulin releasing, pepsin secretion promoting and
 CC adipolytic action. It is used as a reagent for study of
 CC pancreatic function or as a remedy for duodenal ulcers.
 XX

SQ Sequence 17 AA;

Query Match 2.3%; Score 7; DB 4; Length 17;
 Best Local Similarity 100.0%; Pred. No. 9.5;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLQ 234
 |||||
 Db 8 rldrlq 14

RESULT 38

P20383
 ID P20383 standard; peptide; 27 AA.

XX
 AC
 XX

XX 30-NOV-1992 (first entry)

XX Protected heptacosapeptide.

XX Secretin; pancreatic juices; gastric juices.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "p-amethoxybenzyloxycarbonyl-protected"
 FT Modified-site 12
 FT Modified-site 14 /note= "NG-mesitylene sulphonylarginine"
 FT Modified-site 18 /note= "NG-mesitylene sulphonylarginine"
 FT Modified-site 21 /note= "NG-mesitylene sulphonylarginine"
 FT Modified-site 21 /note= "NG-mesitylene sulphonylarginine"

XX JP56158747-A.

XX 07-DEC-1981.

XX 19-OCT-1979; 79JP-0086417.

XX 12-MAY-1980; 80JP-0063174.

XX (NNSH) NIPPON SHINYAKU KK.

XX WPI; 1982-04870E/03 (04870E).

XX Para:methoxy:benzyloxy:carbonyl protected heptacosapeptide - is
 PT intermediate for secretin, which e.g. stimulates pancreatic
 PT juices

XX Claim 1; Page 1; 5pp; Japanese.

XX The sequence given is a heptacosapeptide which can be used as a
 CC precursor for secretin production. Secretin is a digestive tract
 CC enzyme which has physiological actions such as pancreatic juice
 CC secretion-stimulating action and gastric juice secretion-inhibiting

CC action. The heptacosapeptide can be converted to secretin by
 CC treating it with CF3SO3H. This yields large amounts of high purity
 CC secretin in a short time.

XX Sequence 27 AA;

Query Match 2.3%; Score 7; DB 3; Length 27;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLQ 234
 |||||
 Db 18 rldrlq 24

RESULT 39

P20398
 ID P20398 standard; peptide; 27 AA.

XX
 AC P20398;

XX 30-NOV-1992 (first entry)

XX Secretin precursor peptide.

XX Strong acid; digestive canal hormone; pancreas; gastrin; pepsin;
 KW insulin.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "Boc protected"
 FT Modified-site 2 /note= "Boc protected"
 FT Modified-site 3 /note= "Boc protected"
 FT Modified-site 5 /note= "OBoc protected"
 FT Modified-site 7 /note= "Boc protected"
 FT Modified-site 8 /note= "Boc protected"
 FT Modified-site 9 /note= "Boc protected"
 FT Modified-site 11 /note= "OBoc protected"
 FT Modified-site 12 /note= "Boc protected"
 FT Modified-site 12 /note= "PhSO2 ring subst. by 1, 2 or 3 alkyl or
 FT alkoxy gps."

FT Modified-site 14 /note= "PhSO2 ring subst. by 1, 2 or 3 alkyl or
 FT alkoxy gps."

FT Modified-site 15 /note= "OBoc protected"

FT Modified-site 16 /note= "Boc protected"

FT Modified-site 18 /note= "PhSO2 ring subst. by 1, 2 or 3 alkyl or
 FT alkoxy gps."

FT Modified-site 21 /note= "PhSO2 ring subst. by 1, 2 or 3 alkyl or
 FT alkoxy gps."

FT Modified-site 21 /note= "PhSO2 ring subst. by 1, 2 or 3 alkyl or
 FT alkoxy gps."

XX EP47997-A.

XX 24-MAR-1982.

XX 11-SEP-1981; 81EP-0107186.

XX 11-SEP-1980; 80JP-0125262.

XX

PA (EISA) EISAI KK.
 XX Uchiyama M, Sato T, Yoshino H, Tsuchiya Y, Konishi M;
 PI Tsujii M, Hisatake Y, Koiwa A;
 XX WPI; 1982-24409E/13 (24409E).
 XX Heptacosapeptide(s) - useful for high yield conversion to high
 PT purity secretin on strong acid treatment
 XX Claim 1; Page 43; 47pp; English.
 XX The sequence in P20398 is a precursor for the production of
 CC secretin. The peptide sequences given in P20399-402 are peptides which
 CC are useful in the production of this precursor. The precursor is
 CC treated with strong acid in the preparation of secretin. Secretin is
 CC one of the digestive canal hormones and is useful in promotion of
 CC pancreatic external secretin, controlling gastrin-stimulating secretin
 CC of the stomach acid, releasing insulin, stimulating secretin of pepsin
 CC and decomposing fat. It is used as a pancreatic-function examining
 CC agent and a medicine for curing duodenal ulcers etc.
 XX Sequence 27 AA;
 SQ

Query Match 2.3%; Score 7; DB 3; Length 27;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLQ 234
 Db 18 rlqrlq 24
 |||||

RESULT 40
 P30551
 ID P30551 standard; Protein; 27 AA.
 AC P30551;
 XX
 XX 05-AUG-1992 (first entry)
 DE Sequence of 27-Deamidosecretin.
 XX secretin; hormone; pancreatic secretion.
 XX DE3220333-A.
 PN 27-JAN-1983.
 XX 28-MAY-1982; 82DE-3220333.
 XX 04-FEB-1982; 82JP-0016734.
 PR 02-JUN-1981; 81JP-0084603.
 XX (WAKU-) WAKUNAGA YAKUHIN.
 XX Suzuki M, Miyaka T, Miyoshi K, Sumi S, Hasegawa A;
 PI Nishizawa T, Fuwa T;
 XX WPI; 1983-09963K/05.
 DR N-PSDB; N30254.
 XX 27-De-amido:secretin useful for stimulating pancreatic secretin
 PT - prepd. by cultivating E.coli transformed with plasmid contg.
 PT synthetic gene
 XX Claim 1; Page 6-7; 57pp; German.
 XX The inventors claim 27-Deamidosecretin, transformed E. coli
 CC specifically the strain XA35(pLS 58), ATCC 39040, and DNA encoding
 CC 27-Deamidosecretin. 27-Deamidosecretin has similar activity to
 CC the natural hormone secretin, i.e. it stimulates pancreatic

CC secretion.
 XX Sequence 27 AA;
 SQ

Query Match 2.3%; Score 7; DB 4; Length 27;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLQ 234
 Db 18 rlqrlq 24
 |||||

RESULT 41
 P30021
 ID P30021 standard; peptide; 27 AA.
 XX P30021;
 AC
 XX 03-SEP-1992 (first entry)
 DT Synthetic secretin.
 XX Pharmaceutically; deprotection; digestive; hormone; pancreatic;
 KW duodenal ulcer.
 KW Synthetic.
 OS JP58144355-A.
 PN 27-AUG-1983.
 PD 22-FEB-1982; 82JP-0026088.
 PF 22-FEB-1982; 82JP-0026088.
 PR (EISA) EISAI KK.
 XX WPI; 1983-779933/40.
 DR Pharmaceutically active secretin - prepd. by removing protective
 PT gp. from heptacosapeptide
 PT Claim 3; Page 2; 13pp; Japanese.
 XX Secretin, which has hitherto been produced by extraction from
 CC porcine duodenum, may be produced by standard solid phase synthesis.
 CC Secretin is a digestive tract hormone with many useful
 CC pharmaceutical actions such as pancreatic secretion promotion,
 CC gastrin stimulation, gastric acid secretion inhibition, insulin
 CC release, stimulation of pepsin secretion and lipolytic action. It
 CC is useful as a reagent for test on pancreatic and as a remedy for
 CC duodenal ulcers.
 XX Sequence 27 AA;
 SQ

Query Match 2.3%; Score 7; DB 4; Length 27;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLQ 234
 Db 18 rlqrlq 24
 |||||

RESULT 42
 P30049
 ID P30049 standard; peptide; 27 AA.
 XX P30049;
 AC
 XX

DT 03-SEP-1992 (first entry)
XX Intermediate in secretin prodn.
DE
XX Protecting gps.: deprotection; digestive tract; hormone; duodenal;
KW ulcer.
XX
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 1 /label= Boc-His-Boc
FT Modified-site 2 /label= Ser-t-Bu
FT Modified-site 3 /label= Asp-O-t-Bu
FT Modified-site 5 /label= Thr-t-Bu
FT Modified-site 7 /label= Thr-t-Bu
FT Modified-site 8 /label= Ser-t-Bu
FT Modified-site 9 /label= Glu-O-t-Bu
FT Modified-site 11 /label= Ser-t-Bu
FT Modified-site 12 /label= Arg(X)
FT Modified-site 14 /note= "X= alkyl- or alkoxy- benzenesulphonyl"
FT Modified-site 15 /label= Arg(X)
FT Modified-site 16 /note= "X= alkyl- or alkoxy- benzenesulphonyl"
FT Modified-site 18 /label= Arg(X)
FT Modified-site 21 /note= "X= alkyl- or alkoxy- benzenesulphonyl"
FT Modified-site 22 /label= Arg(X)
FT /note= "X= alkyl- or alkoxy- benzenesulphonyl"
XX JP58152853-A.
XX 10-SEP-1983.
XX 05-MAR-1982; 82JP-0034028.
XX 05-MAR-1982; 82JP-0034028.
XX (EISA) EISAI KK.
XX WPI; 1983-791976/42.
XX Heptacosapeptide - obtd. from hepta:peptide and deca:peptide
XX Claim 1; Page 1; 12pp; Japanese.
XX The C-terminal amidated protected peptide may be used to prepare
CC secretin, by deprotection with strong acid. The peptide is an
CC intermediate for the industrial prodn. of secretin. Secretin is a
CC digestive tract hormone. It has pancreatic exocrinogenic, gastrin
CC stimulating, gastric acid secretion inhibiting, insulin releasing,
CC pepsin secretion promoting and adipolytic activities. Secretin may
CC be used as a reagent for the study of pancreatic function and as a
CC remedy for duodenal ulcers.
XX Sequence 27 AA;
XX
XX Query Match 2.3%; Score 7; DB 4; Length 27;
XX Best Local Similarity 100.0%; Pred. No. 14;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 228 RLQRLQ 234
DB 18 rlqrllq 24
RESULT 43
P30014
ID P30014 standard; peptide; 27 AA.
XX
AC P30014;
XX
DT 11-SEP-1992 (first entry)
XX
DE 27-Desamidosecretin.
XX
KW Diagnosis; duodenal ulcer; pancreas.
XX
OS Synthetic.
XX
PN JP57200343-A.
XX
PD 08-DEC-1982.
XX
PF 02-JUN-1981; 81JP-0084603.
XX
PR 02-JUN-1981; 81JP-0084603.
PR 01-JAN-1981; 81JP-0106607.
PR 04-FEB-1982; 82JP-0016734.
PR 02-JUN-1981; 81JP-0084604.
XX
PA (WAKI-) WAKINAGA YAKUHI KK.
XX
XX WPI; 1983-08056K/04 (08056K).
XX
PT 27-Des-amido-secretin prepd. by recombinant DNA techniques -
PT useful as diagnostic agent for pancreatic function or drug for
PT treating duodenal ulcers
XX
PS Claim 1; Page 1; 15pp; Japanese.
XX
XX Prodn. of the peptide comprises chemical synthesis of the peptide
CC expression gene, introduction of the gene into a plasmid capable of
CC growing in a host microorganism, thereby giving a chimeric plasmid
CC which can grow in the microorganism, transformation of the host cell
CC by the plasmid and cultivation of the resultant transformant and
CC recovery of the peptide. The peptide is useful as a diagnostic
CC agent for pancreatic function or as a drug for treatment of duodenal
CC tumour. The peptide is produced by recombinant DNA technique in
CC good yield on large scale with low cost.
XX
SQ Sequence 27 AA;
XX
XX Query Match 2.3%; Score 7; DB 4; Length 27;
XX Best Local Similarity 100.0%; Pred. No. 14;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 228 RLQRLQ 234
DB 18 rlqrllq 24
RESULT 44
P30038
ID P30038 standard; peptide; 27 AA.
XX
AC P30038;
XX
DT 04-SEP-1992 (first entry)
XX
DE Pig Secretin.

XX Porcine; digestive; hormone; pancreatic; duodenal ulcer.
 XX
 XX

OS Sus scrofa.

XX Key Location/Qualifiers
 FT Modified-site 27
 FT /label= Val-X
 FT /note= "X= NH2"

XX JP58152852-A.

XX 10-SEP-1983.

XX 05-MAR-1982; 82JP-0034027.

XX 05-MAR-1982; 82JP-0034027.

XX (EISA) EISAI KK.

XX WPI; 1983-791975/42.

XX Deca:peptide useful as intermediate for secretin - contains
 PT histidine, serine, aspartic acid, glycine, threonine,
 PT phenylalanine, glutamic acid and leucine

XX Disclosure; Page 1; 13pp; Japanese.

XX The peptide, secretin, may be isolated from pigs by standard methods.
 CC Alternatively the peptide may be produced by synthetic intermediates.
 CC Secretin is a digestive tract hormone. It displays pancreatic
 CC exocrinogenic, gastrin stimulating, gastric acid secretion
 CC inhibiting, insulin releasing, pepsin secretion promoting and
 CC adipolytic action.
 CC See also P30039.

XX Sequence 27 AA;

Query Match 2.3%; Score 7; DB 4; Length 27;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 228 RLQRLQ 234
 |||||
 Db 18 rlqlqlq 24

RESULT 45

ID P60647 standard; peptide; 27 AA.

XX AC P60647;

XX 23-JUN-1991 (first entry)

XX Secretin protein sequence.

XX Secretin; hormone.

XX Homo sapiens.

XX WO8605494-A.

XX 25-SEP-1986.

XX 07-MAR-1986; 86WO-SE00099.

XX 11-MAR-1985; 85SE-0001202.

XX (KABI) KABIGEN AB.

XX (SKAN-) SKANDIGEN AB.

XX

PI Carlquist M, Jornvall H, Forssmann W, Thulin L, Johansson C;
 PI Mutt V;

XX WPI; 1986-264936/40.

XX Human intestinal hormone secretin isolated from human duodeni -
 PT useful as diagnostic to determine pancreatic and gall bladder
 PT functions and therapeutically to treat gastro-intestinal
 PT disorders.

XX Claim 1; Page 8; 10pp; English.

XX The sequence encodes the human intestinal hormone, secretin,
 CC which stimulates secretion of water and bicarbonate from the
 CC pancreas. It can be used diagnostically to determine pancreatic
 CC and gall bladder functions, or therapeutically to treat gastro-
 CC intestinal disorders.

XX Sequence 27 AA;

Query Match 2.3%; Score 7; DB 7; Length 27;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 228 RLQRLQ 234
 |||||
 Db 18 rlqlqlq 24

RESULT 46

ID R93024 standard; Protein; 27 AA.

XX AC R93024;

XX 09-AUG-1996 (first entry)

XX Human glucagon degrading enzyme - selectin substrate.

XX Glucagon degrading enzyme; catalyst; cleavage; selectin; human; primer;
 KW vasoactive intestinal peptide; VIP; pancreatic carcinoma cell line; PCR;
 KW amplification; polymerase chain reaction; probe; expression vector;
 KW eukaryote; SV40 promoter; COS-7.

XX Synthetic.

XX Key Location/Qualifiers
 FT Cleavage-site 14..15
 FT Modified-site 27

FT /note= "contains C-terminal amide group"

XX JP08023972-A.

XX 30-JAN-1996.

XX 19-JUL-1994; 94JP-0187936.

XX 19-JUL-1994; 94JP-0187936.

XX (SUNR) SUNTORY LTD.

XX WPI; 1996-133414/14.

XX New glucagon decomposing enzyme, and DNA encoding it - for
 PT specifically cleaving glucagon and vasoactive intestinal peptide, in
 PT the prevention and treatment of diseases caused by excess glucagon
 PT and Vip

XX Claim 1; Page 2; 18pp; Japanese.

XX A novel gene encoding a glucagon degrading enzyme (GDE; T11575) was
 CC isolated from a human pancreatic carcinoma cell line HPC-Yo cDNA library.

CC The enzyme has a mol. wt. 83 kD, a pH optimum of 6.8 and catalyses the
CC cleavage of glucagon, vasoactive intestinal peptide and selectin
CC (R93022-4). The gene encoding the enzyme was isolated by screening
CC the library with an anti-GDE peptide antibody, amplifying the inserts
CC with the primers T18903-4 and probing the fragments with the probe
CC T18905. This screening resulted in the full length clone designated
CC lambda GDE4-2. The coding region of the clone was subsequently PCR
CC amplified by the primers T11576-7 and inserted into the eukaryotic
CC expression vector pKPCR under control of the SV40 promoter for production
CC of the protein in COS-7 cells. The protein is useful in preventing and
CC treating diseases characterised by an excess of glucagon or vasoactive
CC intestinal peptide.

XX
SQ Sequence 27 AA;

Query Match 2.3%; Score 7; DB 17; Length 27;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLIQ 234
Db |||||
18 rlqrlilq 24

RESULT 47

W11676
ID W11676 standard; Peptide; 27 AA.

XX
AC W11676;

XX
DT 11-JAN-1999 (first entry)

XX Secretin-derived target peptide.

XX Calmodulin; green fluorescent protein; GFP; cameleon;
KW Fluorescence resonance energy transfer; FRET; calcium; sensor;
KW analysis; assay; secretin.

XX Synthetic.

XX
PN WO9840477-A1.

XX
PD 17-SEP-1998.

XX
PF 13-MAR-1998; 98WO-US04978.

XX
PR 27-AUG-1997; 97US-0919143.

XX
PR 14-MAR-1997; 97US-0818252.

XX
PR 14-MAR-1997; 97US-0818253.

XX (REGC) UNIV CALIFORNIA.

XX
PI Miyawaki A, Tsien RY;

XX
DR WPI; 1998-520809/44.

XX
PT New fluorescent protein sensors for detection of analytes -
PT comprises a binding protein moiety having an analyte binding region
PT and bound donor and acceptor fluorescent protein moieties

XX
PS Disclosure; Page 21; 108pp; English.

XX This peptide represents a target moiety from secretin that is
CC recognised by calmodulin. The invention provides fluorescent
CC indicators and methods for using them to determine the
CC concentration of an analyte, such as calcium ion, in vitro and in
CC vivo. Fluorescent indicators include a binding protein moiety
CC (e.g. calmodulin) and donor and acceptor fluorescent protein
CC moieties, preferably derived from Aequorea green fluorescent
CC protein (see W11645-48). The binding protein preferably binds
CC target peptides (see W11649-79) in addition to the analyte. The
CC target peptide moieties can be modified to enhance the response of

CC the fluorescent indicator to the analyte.

XX
SQ Sequence 27 AA;

Query Match 2.3%; Score 7; DB 19; Length 27;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLIQ 234
Db |||||
18 rlqrlilq 24

RESULT 48

W37793
ID W37793 standard; peptide; 27 AA.

XX
AC W37793;

XX
DT 28-JUL-1998 (first entry)

XX Porcine secretin peptide.

XX Porcine secretin; vasoactive intestinal polypeptide-1 receptor;
KW VIP-1 receptor; peptidic ligand; VIP-2 receptor; agonist; antagonist;
KW bronchoconstrictive disorder; asthma; tumour; stroke; cancer;
KW chronic obstructive pulmonary disease; myocardial infarction;
KW gastroenterological disease; anti-inflammatory; cell growth;
KW organ transplantation; cancer.

XX Sus scrofa.

XX
FH Key Location/Qualifiers
FT Modified-site 27
FT /label= val
FT /note= "amidated"

XX
PN WO9802453-A2.

XX
PD 22-JAN-1998.

XX
PF 15-JUL-1997; 97WO-BE00084.

XX
PR 19-SEP-1996; 96EP-0870121.

XX
PR 15-JUL-1996; 96EP-0870092.

XX (ULBR) UNIV LIBRE BRUXELLES.

XX Gourlet P, Robberecht P, Vandermeers A, Woelbroeck M;

XX
DR WPI; 1998-110523/10.

XX New ligands for vasoactive intestinal peptide receptor - is useful
PT for treating VIP-related disorders, e.g. asthma, tumours, myocardial
PT infarction, stroke, inflammation or auto-immune disease

XX
PS Example 1; Page 18; 38pp; English.

XX This is the amino acid sequence of a porcine secretin, used as a
CC comparison for the vasoactive intestinal polypeptide (VIP) in the
CC method of the invention. VIP has two distinct receptors with seven
CC transmembrane helices named VIP-1 and VIP-2. The method of the
CC invention involves the development of peptidic ligands that can be
CC used in the treatment of bronchoconstrictive disorders, e.g. asthma,
CC chronic obstructive pulmonary disease (COPD), tumours, myocardial
CC infarctions, strokes, the regeneration of nerves as in post-traumatic
CC injury, as anti-inflammatory and anti-oxidant agent, to increase cell
CC growth, as immuno-modulation agent in the treatment of auto-immune
CC diseases and for reducing side effects in organ transplantation. They
CC can also be used for detection and diagnosis, e.g. for the
CC identification of specific cancers such as breast and prostate cancers,
CC lung cancers, ovarian cancers and colon cancers. The ligands can also

CC be used for the identification of other ligands of the VIP1 receptor.

XX Sequence 27 AA;

Query Match 2.3%; Score 7; DB 19; Length 27;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 228 RLRLQLQ 234
| | | | | | |
Db 18 RLRLQLQ 24

RESULT 49

W37794
ID W37794 standard; peptide; 27 AA.

XX AC W37794;

XX DT 28-JUL-1998 (first entry)

XX DE Rabbit secretin peptide.

XX KW Rabbit secretin; vasoactive intestinal polypeptide-1 receptor;
KW VIP-1 receptor; peptidic ligand; VIP-2 receptor; agonist; antagonist;
KW bronchoconstrictive disorder; asthma; tumour; stroke; cancer;
KW chronic obstructive pulmonary disease; myocardial infarction;
KW gastroenterological disease; anti-inflammatory; cell growth;
KW organ transplantation; cancer.

XX OS Oryctolagus cuniculus.

XX FH Key Location/Qualifiers
XX FT Modified-site 27
XX FT /label= Leu
XX FT /note= "amidated"

XX PN W09802453-A2.

XX PD 22-JAN-1998.

XX PF 15-JUL-1997; 97WO-BE00084.

XX PR 19-SEP-1996; 96EP-0870121.

XX PR 15-JUL-1996; 96EP-0870092.

XX PA (ULBR) UNIV LIBRE BRUXELLES.

XX PI Gourlet P, Robberecht P, Vandermeers A, Woelbroeck M;

XX DR WPI; 1998-110523/10.

XX PT New ligands for vasoactive intestinal peptide receptor - is useful
XX PT for treating VIP-related disorders, e.g. asthma, tumours, myocardial
XX PT infarction, stroke, inflammation or auto-immune disease

XX PS Example 1; Page 18; 38pp; English.

XX CC This is the amino acid sequence of a rabbit secretin, used as a
XX CC comparison for the vasoactive intestinal polypeptide (VIP) in the
XX CC method of the invention. VIP has two distinct receptors with seven
XX CC transmembrane helices named VIP-1 and VIP-2. The method of the
XX CC invention involves the development of peptidic ligands that can be
XX CC used in the treatment of bronchoconstrictive disorders, e.g. asthma,
XX CC chronic obstructive pulmonary disease (COPD), tumours, myocardial
XX CC infarctions, strokes, the regeneration of nerves as in post-traumatic
XX CC injury, as anti-inflammatory and anti-oxidant agent, to increase cell
XX CC growth, as immuno-modulation agent in the treatment of auto-immune
XX CC diseases and for reducing side effects in organ transplantation. They
XX CC can also be used for detection and diagnosis, e.g. for the
XX CC identification of specific cancers such as breast and prostate cancers,
XX CC lung cancers, ovarian cancers and colon cancers. The ligands can also

CC be used for the identification of other ligands of the VIP1 receptor.

XX Sequence 27 AA;

Query Match 2.3%; Score 7; DB 19; Length 27;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 228 RLRLQLQ 234
| | | | | | |
Db 18 RLRLQLQ 24

RESULT 50

W37795
ID W37795 standard; peptide; 27 AA.

XX AC W37795;

XX DT 28-JUL-1998 (first entry)

XX DE Porcine secretin peptide variant 1.

XX KW Porcine secretin; vasoactive intestinal polypeptide-1 receptor;
KW VIP-1 receptor; peptidic ligand; VIP-2 receptor; agonist; antagonist;
KW bronchoconstrictive disorder; asthma; tumour; stroke; cancer;
KW chronic obstructive pulmonary disease; myocardial infarction;
KW gastroenterological disease; anti-inflammatory; cell growth;
KW organ transplantation; cancer.

XX OS Sus scrofa.

XX FH Key Location/Qualifiers
XX FT Modified-site 27
XX FT /label= Val
XX FT /note= "amidated"

XX PN W09802453-A2.

XX PD 22-JAN-1998.

XX PF 15-JUL-1997; 97WO-BE00084.

XX PR 19-SEP-1996; 96EP-0870121.

XX PR 15-JUL-1996; 96EP-0870092.

XX PA (ULBR) UNIV LIBRE BRUXELLES.

XX PI Gourlet P, Robberecht P, Vandermeers A, Woelbroeck M;

XX DR WPI; 1998-110523/10.

XX PT New ligands for vasoactive intestinal peptide receptor - is useful
XX PT for treating VIP-related disorders, e.g. asthma, tumours, myocardial
XX PT infarction, stroke, inflammation or auto-immune disease

XX PS Example 1; Page 18; 38pp; English.

XX CC This is the amino acid sequence of the porcine secretin variant 1,
XX CC used as a comparison for the vasoactive intestinal polypeptide (VIP)
XX CC in the method of the invention. VIP has two distinct receptors with
XX CC seven transmembrane helices named VIP-1 and VIP-2. The method of the
XX CC invention involves the development of peptidic ligands that can be
XX CC used in the treatment of bronchoconstrictive disorders, e.g. asthma,
XX CC chronic obstructive pulmonary disease (COPD), tumours, myocardial
XX CC infarctions, strokes, the regeneration of nerves as in post-traumatic
XX CC injury, as anti-inflammatory and anti-oxidant agent, to increase cell
XX CC growth, as immuno-modulation agent in the treatment of auto-immune
XX CC diseases and for reducing side effects in organ transplantation. They
XX CC can also be used for detection and diagnosis, e.g. for the
XX CC identification of specific cancers such as breast and prostate cancers,
XX CC lung cancers, ovarian cancers and colon cancers. The ligands can also

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:55:16 ; Search time 27.49 Seconds
(without alignments)
741.004 Million cell updates/sec

Title: US-09-518-931-2
Perfect score: 300
Sequence: 1 MRALEGPGLSLLCLVLALPA.....RVARMPLGERSVREFLPVH 300

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 195891 seqs, 67900655 residues

Word size : 6

Total number of hits satisfying chosen parameters: 1995

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : PIR_66:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	3.0	327	2 H83483	probable transmembr
2	8	2.7	179	1 KGR7	gamma-casein precu
3	8	2.7	181	2 T49104	hypothetical prote
4	8	2.7	184	2 I49685	gamma-casein precu
5	8	2.7	241	2 T05479	hypothetical prote
6	8	2.7	341	2 T27654	hypothetical prote
7	8	2.7	440	2 T50912	hypothetical prote
8	8	2.7	494	1 S60028	ferredoxin--NADP+
9	8	2.7	530	2 B64905	sugar kinase homol
10	8	2.7	531	1 S54098	1-phosphatidylinos
11	8	2.7	644	2 JC5119	anti-mullerian hor
12	8	2.7	781	2 T49472	hormone-sensitive
13	8	2.7	1172	1 TSHUP2	thrombospondin 2 p
14	7	2.3	26	1 B57082	secretin - guinea
15	7	2.3	27	1 S07443	secretin - human
16	7	2.3	27	1 SEBO	secretin - bovine
17	7	2.3	27	1 SESH	secretin - sheep
18	7	2.3	27	2 A27267	secretin - dog
19	7	2.3	27	2 C60415	secretin - rabbit
20	7	2.3	74	2 S13515	retinoic acid rece
21	7	2.3	89	2 S13517	retinoic acid rece
22	7	2.3	106	2 G82729	hypothetical prote
23	7	2.3	112	2 G72502	hypothetical prote
24	7	2.3	113	1 IMECE1	colicin E1 immunit
25	7	2.3	113	2 I64785	imm protein - Esch
26	7	2.3	113	2 S11532	colicin E1 immunit
27	7	2.3	114	2 S44660	ZK353.5 protein -
28	7	2.3	118	2 S27476	hypothetical prote
29	7	2.3	131	1 SEPG	secretin precursor

30	7	2.3	133	2 JC2202	secretin precursor
31	7	2.3	134	2 A40959	secretin precursor
32	7	2.3	140	2 T27059	hypothetical prote
33	7	2.3	140	2 C72705	hypothetical prote
34	7	2.3	144	1 TVVPBD	small T antigen -
35	7	2.3	150	2 S34380	hypothetical prote
36	7	2.3	164	2 T09900	hypothetical prote
37	7	2.3	166	2 T25118	hypothetical prote
38	7	2.3	168	2 E83573	hypothetical prote
39	7	2.3	182	2 T16519	hypothetical prote
40	7	2.3	182	2 S69469	hypothetical prote
41	7	2.3	198	2 S48994	hypothetical prote
42	7	2.3	201	2 G69906	conserved hypothet
43	7	2.3	208	2 C75041	orotidine-5'-monop
44	7	2.3	210	2 G83082	hypothetical prote
45	7	2.3	215	2 S34163	homeotic protein H
46	7	2.3	229	2 JC7219	nuclear protein SR
47	7	2.3	230	2 T24789	hypothetical prote
48	7	2.3	235	2 A48383	alpha s2-casein -
49	7	2.3	241	2 D71120	hypothetical prote
50	7	2.3	272	2 T05498	hypothetical prote
51	7	2.3	274	2 C71685	hypothetical prote
52	7	2.3	276	2 T51685	myb-related transc
53	7	2.3	277	2 G71976	type II DNA modifi
54	7	2.3	277	2 D64531	type II restrictio
55	7	2.3	282	2 S74617	prohibitin phb - S
56	7	2.3	285	2 T06627	hypothetical prote
57	7	2.3	292	2 JE0233	troponin-I - scal
58	7	2.3	292	2 T36914	probable integral
59	7	2.3	296	2 T02573	CDP-diacylglycerol
60	7	2.3	303	2 B70554	hypothetical prote
61	7	2.3	310	2 T31799	hypothetical prote
62	7	2.3	312	2 T08985	hypothetical prote
63	7	2.3	323	2 S72915	hypothetical prote
64	7	2.3	334	2 JC6082	proximal sequence
65	7	2.3	334	2 G02375	SNAP45 subunit - h
66	7	2.3	335	2 A72065	GTP binding protei
67	7	2.3	338	2 T26399	hypothetical prote
68	7	2.3	338	2 H72115	cationic amino aci
69	7	2.3	343	2 G81602	GPPI/OBG family pr
70	7	2.3	346	2 T32585	hypothetical prote
71	7	2.3	349	2 T41394	hypothetical serin
72	7	2.3	353	2 A43697	homeotic protein z
73	7	2.3	354	2 B55733	G protein-coupled
74	7	2.3	354	2 S39406	homeotic protein o
75	7	2.3	355	2 S35345	otxl protein - mou
76	7	2.3	355	2 I56547	homeodomain protei
77	7	2.3	364	1 JE0292	fungal stress prot
78	7	2.3	364	2 S43574	COS95.3 protein (c
79	7	2.3	368	2 T51200	hypothetical prote
80	7	2.3	369	2 T17267	hypothetical prote
81	7	2.3	369	2 T01196	transcription fact
82	7	2.3	381	2 T13666	NADH dehydrogenase
83	7	2.3	389	2 G83317	conserved hypothet
84	7	2.3	391	2 S61704	probable transcrip
85	7	2.3	392	2 T43490	hypothetical prote
86	7	2.3	393	2 E72549	probable acyl-CoA
87	7	2.3	406	2 G70410	glutamyl tRNA redu
88	7	2.3	406	2 S38170	SRP40 protein - ye
89	7	2.3	406	2 T30748	hypothetical prote
90	7	2.3	407	2 E71665	bicyclomycin resis
91	7	2.3	411	2 S34164	homeotic protein H
92	7	2.3	412	2 A70416	ATP-dependent clp
93	7	2.3	414	2 T47281	hypothetical prote
94	7	2.3	417	2 S47539	homeotic protein H
95	7	2.3	421	2 T38309	probable cell wall
96	7	2.3	423	2 B72528	phytoene nucleolar
97	7	2.3	425	2 T10702	phytoene synthase
98	7	2.3	429	2 H71539	probable dihydrol
99	7	2.3	432	2 T35527	hypothetical prote
100	7	2.3	433	2 T36609	probable secreted
101	7	2.3	434	2 S29612	pectate lyase (EC
102	7	2.3	444	2 T22412	hypothetical prote

103	7	2.3	453	2	H70333	conserved hypothet	176	7	2.3	1217	2	S52714	sericinB - silkwo
104	7	2.3	455	2	B83213	probable MFS trans	177	7	2.3	1279	2	T18312	hypothetical prote
105	7	2.3	457	2	T21063	hypothetical prote	178	7	2.3	1288	2	JE0363	mitogen-activated
106	7	2.3	458	2	T02804	hypothetical prote	179	7	2.3	1298	1	EDBE75	immediate-early pr
107	7	2.3	468	2	S70297	SPS2 protein homol	180	7	2.3	1324	2	T01508	mismatch repair en
108	7	2.3	469	2	A35789	glutamate-1-semial	181	7	2.3	1339	2	A55301	hypothetical prote
109	7	2.3	472	2	T22188	hypothetical prote	182	7	2.3	1367	2	S51959	1,3-beta-D-glucan-
110	7	2.3	479	2	G83071	outer membrane pro	183	7	2.3	1417	2	A57570	Bloom's syndrome r
111	7	2.3	486	2	T35445	probable integral	184	7	2.3	1440	2	T44872	probable integral
112	7	2.3	487	1	QQEGRS	ygjE protein - Esc	185	7	2.3	1446	2	A45344	immediate-early pr
113	7	2.3	494	1	B45738	alpha-amyrase (EC	186	7	2.3	1460	1	EDBEIF	immediate-early pr
114	7	2.3	495	1	A45738	alpha-amyrase (EC	187	7	2.3	1469	2	T09219	basal transcriptio
115	7	2.3	500	2	T14826	transcription init	188	7	2.3	1487	2	T02850	hypothetical prote
116	7	2.3	504	1	T1WL51	L1 protein - human	189	7	2.3	1548	2	T04456	hypothetical prote
117	7	2.3	508	2	C82138	conserved hypothet	190	7	2.3	1681	2	A55138	sodium channel mna
118	7	2.3	509	2	S08059	alpha-globulin typ	191	7	2.3	2185	1	S60200	acetyl-CoA carboxy
119	7	2.3	509	2	T45868	hypothetical prote	192	7	2.3	2187	2	T30826	nascent polypeptid
120	7	2.3	513	2	T31115	histidine kinase h	193	7	2.3	2232	2	T34434	hypothetical prote
121	7	2.3	517	2	G83311	probable transcript	194	7	2.3	2422	2	T12687	ALR protein homolo
122	7	2.3	519	2	T16712	hypothetical prote	195	7	2.3	2560	1	I40457	peptide synthetase
123	7	2.3	526	2	JE0131	glucose-1-phosphat	196	7	2.3	3910	2	A44264	trithorax homolog
124	7	2.3	546	2	T47018	hypothetical prote	197	7	2.3	3968	2	A44265	trithorax homolog
125	7	2.3	548	2	S59133	ETS2 repressor fac	198	7	2.3	4563	1	LPHUB	apolipoprotein B-1
126	7	2.3	554	1	TVVPBF	large T antigen -	199	6	2.0	25	2	PC4437	hydrogenase (EC 1.
127	7	2.3	557	2	T41588	probable amino-aci	200	6	2.0	28	2	T14905	hypothetical prote
128	7	2.3	559	1	RWBYS1	glycophospholipid-	201	6	2.0	39	2	F82767	hypothetical prote
129	7	2.3	567	2	T08405	hypothetical prote	202	6	2.0	47	1	W0BP57	gene 0.5 protein -
130	7	2.3	568	2	S42074	aspartic proteinas	203	6	2.0	47	2	S78708	protein YBR126W-a
131	7	2.3	600	2	S76764	hypothetical prote	204	6	2.0	56	2	E72718	hypothetical prote
132	7	2.3	600	2	S07638	spore coat protein	205	6	2.0	61	2	S05999	xylokinnase (EC 2
133	7	2.3	601	2	E71435	probable triacylgl	206	6	2.0	70	2	H72168	A43R protein - var
134	7	2.3	606	2	JC5604	ABC-transporting p	207	6	2.0	72	1	OSB08	cytochrome-c oxida
135	7	2.3	623	2	T38000	anthranilate synth	208	6	2.0	73	2	F83414	hypothetical prote
136	7	2.3	630	2	T38023	probable transcript	209	6	2.0	76	2	S37191	hypothetical prote
137	7	2.3	636	2	S63131	probable membrane	210	6	2.0	80	2	T36813	hypothetical prote
138	7	2.3	640	1	QQBEU2	UL35 protein - hum	211	6	2.0	80	2	T27603	hypothetical prote
139	7	2.3	641	2	T03095	homeoprotein Sail	212	6	2.0	81	2	S47426	envelope protein s
140	7	2.3	642	2	S71437	sulfite reductase	213	6	2.0	81	2	C60076	hypothetical prote
141	7	2.3	642	2	T48455	sulfite reductase	214	6	2.0	82	2	S24280	hypothetical prote
142	7	2.3	644	2	S39356	transcription fact	215	6	2.0	82	2	C60007	hypothetical prote
143	7	2.3	659	2	S67175	probable membrane	216	6	2.0	82	2	B64791	YbdJ protein - Esc
144	7	2.3	671	2	S51599	Om(2D) protein - f	217	6	2.0	84	2	T34109	hypothetical prote
145	7	2.3	697	2	T07111	MAR binding filame	218	6	2.0	88	2	D72803	gp30 protein - Myc
146	7	2.3	708	2	S52317	quinohemoprotein e	219	6	2.0	90	2	T10332	FP protein - Orgy1
147	7	2.3	731	2	A83160	conserved hypothet	220	6	2.0	92	2	I55688	seminal vesicle pr
148	7	2.3	733	2	T28145	RING3 kinase - chi	221	6	2.0	92	2	S23527	embryonic abundant
149	7	2.3	737	2	S47857	basic protein, cyt	222	6	2.0	95	2	D64389	hypothetical prote
150	7	2.3	741	2	T46488	hypothetical prote	223	6	2.0	96	2	T08761	hypothetical prote
151	7	2.3	743	2	T02828	conserved hypothet	224	6	2.0	98	1	A27510	ferredoxin 2[4Fe-4
152	7	2.3	744	2	S65669	biotin sulfoxide r	225	6	2.0	99	2	B83516	conserved hypothet
153	7	2.3	754	2	A56619	female sterile hom	226	6	2.0	100	2	T11037	hypothetical prote
154	7	2.3	792	2	A83303	phenylalanyl-tRNA	227	6	2.0	102	2	E70073	hypothetical prote
155	7	2.3	801	1	A44165	1-phosphatidylinos	228	6	2.0	102	2	T34642	hypothetical prote
156	7	2.3	806	2	B49338	hupU protein - Rho	229	6	2.0	102	2	S67179	probable membrane
157	7	2.3	837	2	H72802	minor tail subunit	230	6	2.0	104	2	S18821	cytochrome c552 pr
158	7	2.3	843	2	S35527	DNA-directed DNA p	231	6	2.0	106	2	T14819	dehydrin DHN2 - so
159	7	2.3	853	2	A71339	probable outer mem	232	6	2.0	106	2	G72632	hypothetical prote
160	7	2.3	864	2	A49070	ecdysone-inducible	233	6	2.0	107	2	E70569	hypothetical prote
161	7	2.3	892	2	T09193	ataxin 7 - human	234	6	2.0	108	2	C64490	hypothetical prote
162	7	2.3	914	2	S18942	hypothetical prote	235	6	2.0	108	2	H83303	hypothetical prote
163	7	2.3	928	2	T20035	hypothetical prote	236	6	2.0	109	2	D65061	hypothetical prote
164	7	2.3	945	2	JN0901	endopeptidase Clp	237	6	2.0	109	2	T12208	probable cytochrom
165	7	2.3	963	2	A55926	DNA binding protei	238	6	2.0	110	2	T08921	hypothetical prote
166	7	2.3	971	2	I39166	cellular apoptosis	239	6	2.0	111	2	JH0183	Sc4 protein - brac
167	7	2.3	1032	2	T34433	hypothetical prote	240	6	2.0	112	2	T47326	hypothetical prote
168	7	2.3	1065	2	S19482	hypothetical prote	241	6	2.0	113	2	A83631	conserved hypothet
169	7	2.3	1070	2	T06733	kinesin homolog F2	242	6	2.0	113	2	D83489	hypothetical prote
170	7	2.3	1070	2	T00767	hypothetical prote	243	6	2.0	117	2	A45509	desiccation-relate
171	7	2.3	1075	2	S48992	flocculation prote	244	6	2.0	117	2	S73891	MG129 homolog A65-
172	7	2.3	1093	2	I38533	AF17 protein - hum	245	6	2.0	119	2	S28133	gas vesicle protei
173	7	2.3	1135	2	T30561	scythe protein - A	246	6	2.0	120	2	S71563	drought-induced pr
174	7	2.3	1150	2	S58775	mypl protein - smu	247	6	2.0	120	2	S65067	dehydrin DHN1 - po
175	7	2.3	1198	2	T20262	hypothetical prote	248	6	2.0	120	2	JQ1740	hypothetical 12.7k

249	6	2.0	120	2	T18015	hypothetical prote	322	6	2.0	155	2	T35313	hypothetical prote
250	6	2.0	121	2	F83373	hypothetical prote	323	6	2.0	155	2	S14239	hypothetical prote
251	6	2.0	122	2	S29540	seminal vesicle se	324	6	2.0	157	2	T10489	dehydrin 1 - Comme
252	6	2.0	122	2	S39457	pollen allergen ph	325	6	2.0	158	2	T31645	hypothetical prote
253	6	2.0	123	2	T22970	hypothetical prote	326	6	2.0	159	2	A75529	hypothetical prote
254	6	2.0	124	2	T24876	hypothetical prote	327	6	2.0	159	2	T15671	hypothetical prote
255	6	2.0	124	2	F72615	hypothetical prote	328	6	2.0	159	2	T36734	probable tetra-fam1
256	6	2.0	124	2	C83137	hypothetical prote	329	6	2.0	161	2	S05547	dehydrin 17 - barl
257	6	2.0	125	2	T23756	hypothetical prote	330	6	2.0	161	2	T47390	hypothetical prote
258	6	2.0	126	2	S74990	hypothetical prote	331	6	2.0	162	2	S38300	protein kinase C i
259	6	2.0	127	1	KNMUHX	dehydrin-like prot	332	6	2.0	163	2	S01152	RAB21 protein - ri
260	6	2.0	127	2	G81688	hypothetical prote	333	6	2.0	163	2	G70827	hypothetical prote
261	6	2.0	128	2	S57955	thrombospondin 2 -	334	6	2.0	164	1	WZPSLP	lactoylglytathione
262	6	2.0	128	2	E72340	hypothetical prote	335	6	2.0	164	2	S11846	rab16B protein - r
263	6	2.0	128	2	T06028	hypothetical prote	336	6	2.0	164	2	S11847	rab16C protein - r
264	6	2.0	128	2	S74258	probable membrane	337	6	2.0	164	2	D70130	conserved hypothet
265	6	2.0	129	2	T46300	hypothetical prote	338	6	2.0	165	2	T03392	probable dehydrin
266	6	2.0	129	2	T36732	hypothetical prote	339	6	2.0	165	2	F69208	hypothetical prote
267	6	2.0	129	2	H72627	hypothetical prote	340	6	2.0	165	2	S68776	myogenin - rainbow
268	6	2.0	129	2	T05474	hypothetical prote	341	6	2.0	166	1	C64687	biopolymer transpo
269	6	2.0	130	2	S12308	TAS14 protein - to	342	6	2.0	166	2	T07089	dehydrin - soybean
270	6	2.0	131	2	A81312	hypothetical prote	343	6	2.0	166	2	S89692	hypothetical prote
271	6	2.0	131	2	S33375	homeotic protein A	344	6	2.0	166	2	F81707	hypothetical prote
272	6	2.0	132	2	G83132	hypothetical prote	345	6	2.0	166	2	S07044	gene B protein - p
273	6	2.0	132	2	T08676	hypothetical prote	346	6	2.0	166	2	T08208	nonstructural prot
274	6	2.0	133	2	D70474	hypothetical prote	347	6	2.0	167	2	F81732	ribosomal protein
275	6	2.0	133	2	E72763	immunoglobulin lam	348	6	2.0	167	2	A71470	probable l9 riboso
276	6	2.0	133	2	E72763	T-cell receptor be	349	6	2.0	167	2	S05545	dehydrin 3 - maize
277	6	2.0	135	2	J00472	dehydrin - upland	350	6	2.0	168	2	A39316	RAB-17 protein - m
278	6	2.0	135	2	T09876	T-cell receptor be	351	6	2.0	168	2	S08633	RAB-17 protein - m
279	6	2.0	136	2	J00473	transformer-2 sex-	352	6	2.0	168	2	T17592	hypothetical prote
280	6	2.0	136	2	D35846	hypothetical prote	353	6	2.0	168	2	S16800	hypothetical prote
281	6	2.0	137	2	S35752	dehydrin 7 - barle	354	6	2.0	169	2	D70906	hypothetical prote
282	6	2.0	137	2	S37195	insulin-like pepti	355	6	2.0	169	2	G83130	hypothetical prote
283	6	2.0	138	2	S38299	protein kinase C i	356	6	2.0	169	2	A81188	spermidine acetyl
284	6	2.0	138	2	A43957	insulin-like pepti	357	6	2.0	170	2	E75358	hypothetical prote
285	6	2.0	139	2	S05544	dehydrin 8 - barle	358	6	2.0	170	2	T49033	zinc finger-like p
286	6	2.0	139	2	T05715	dehydrin - barley	359	6	2.0	171	2	B70408	conserved hypothet
287	6	2.0	139	2	C70660	hypothetical prote	360	6	2.0	171	2	H81914	hypothetical prote
288	6	2.0	139	2	T30863	stbA protein - Sal	361	6	2.0	172	1	QBED7	HHLF4 protein - hu
289	6	2.0	140	2	C71326	probable V-type AT	362	6	2.0	172	2	E83301	hypothetical prote
290	6	2.0	141	2	B83079	hypothetical prote	363	6	2.0	172	2	T36107	probable serine/ar
291	6	2.0	142	2	C75564	hypothetical prote	364	6	2.0	173	2	T48097	glutathione peroxi
292	6	2.0	142	2	J01865	hypothetical 15.1K	365	6	2.0	174	2	T15176	hypothetical prote
293	6	2.0	143	2	S05543	dehydrin 9 - barle	366	6	2.0	174	2	T37014	probable transposa
294	6	2.0	143	2	T02321	hypothetical prote	367	6	2.0	174	2	C59268	conserved hypothet
295	6	2.0	144	2	T44390	ribosomal protein	368	6	2.0	175	2	A71227	probable cob(l)ala
296	6	2.0	144	2	B83069	hypothetical prote	369	6	2.0	175	2	S77306	hypothetical prote
297	6	2.0	145	2	S04042	embryonic abundant	370	6	2.0	175	2	H69311	conserved hypothet
298	6	2.0	145	2	D71829	biopolymer transpo	371	6	2.0	175	2	D70824	hypothetical prote
299	6	2.0	145	2	A47688	hemagglutinating s	372	6	2.0	175	2	D83338	hypothetical prote
300	6	2.0	145	2	C75590	hypothetical prote	373	6	2.0	176	1	A46606	platelet glycoprot
301	6	2.0	148	2	S02509	nifz protein - Kle	374	6	2.0	176	2	T04526	hypothetical prote
302	6	2.0	148	2	G70890	probable repressor	375	6	2.0	178	2	T25432	hypothetical prote
303	6	2.0	149	2	S19130	rab15 protein - wh	376	6	2.0	180	2	T27193	beta-lactoglobulin
304	6	2.0	149	2	C75598	response regulator	377	6	2.0	180	1	LGST	heat shock protein
305	6	2.0	149	2	T23179	hypothetical prote	378	6	2.0	180	1	LGSH	finger protein XFG
306	6	2.0	149	2	D72526	hypothetical prote	379	6	2.0	180	2	S56943	14-3-3 protein - f
307	6	2.0	150	2	C82994	conserved hypothet	380	6	2.0	180	2	S56026	WD repeat protein
308	6	2.0	150	2	T01827	hypothetical prote	381	6	2.0	181	2	T12088	hypothetical prote
309	6	2.0	150	2	J50297	infection structur	382	6	2.0	181	3	T41517	beta-lactoglobulin
310	6	2.0	151	1	GGICEA	globin CTT-IIia -	383	6	2.0	182	2	T23205	hypothetical prote
311	6	2.0	151	2	J02282	negatively phytoch	384	6	2.0	183	2	B83441	potassium-transpor
312	6	2.0	151	2	S11848	rab16D protein - r	385	6	2.0	184	2	S12095	embryonic abundant
313	6	2.0	151	2	T20099	hypothetical prote	386	6	2.0	186	1	WMR219	19K globulin precu
314	6	2.0	151	2	F83095	conserved hypothet	387	6	2.0	186	2	JC4784	alpha-globulin pre
315	6	2.0	151	2	A82978	hypothetical prote	388	6	2.0	186	2	S28021	rab18 protein - Ar
316	6	2.0	152	2	S35203	serine proteinase	389	6	2.0	186	2	C81727	translation initia
317	6	2.0	153	2	S12206	hypothetical prote	390	6	2.0	186	2	C75095	hypothetical prote
318	6	2.0	154	2	T14727	dehydrin DHN1 - so	391	6	2.0	186	2	D75005	cdp-diacylglycerol
319	6	2.0	154	2	T36349	probable regulator	392	6	2.0	186	2	T21961	hypothetical prote
320	6	2.0	155	1	T12M1	trypsin/factor XII	393	6	2.0	186	2	T32505	hypothetical prote
321	6	2.0	155	2	S43775	desiccation-relate	394	6	2.0	186	2	A82116	conserved hypothet

395	6	2.0	187	2	C70213	hypothetical prote	468	2.0	215	2	S29595	calmodulin-related
396	6	2.0	188	2	S59915	14-3-3 protein iso	469	2.0	216	2	T29039	hypothetical prote
397	6	2.0	188	2	A64332	probable phosphono	470	2.0	216	2	E82326	MSHA biogenesis pr
398	6	2.0	189	2	S43558	membrane protein B	471	2.0	216	2	F70073	hypothetical prote
399	6	2.0	189	2	T27809	hypothetical prote	472	2.0	216	2	S50766	dehydrin-related p
400	6	2.0	189	2	T34716	hypothetical prote	473	2.0	218	1	VCVX11	coat protein - cuc
401	6	2.0	189	2	T31525	hypothetical prote	474	2.0	218	2	D71392	peptide methionine
402	6	2.0	190	2	E81109	lipote-protein li	475	2.0	218	2	T49885	hypothetical prote
403	6	2.0	190	2	E81907	probable lipote-p	476	2.0	218	2	G75290	hypothetical prote
404	6	2.0	190	2	G83377	hypothetical prote	477	2.0	218	2	S74867	protein-L-isoaspar
405	6	2.0	190	2	T35951	probable SIR2 faml	478	2.0	218	2	H82539	probable lipote-p
406	6	2.0	191	2	D64797	lipote-protein l	479	2.0	219	2	T38809	14-3-3 protein hom
407	6	2.0	191	2	JS0582	hypothetical 22K p	480	2.0	220	2	S20581	antistatin - Hydra
408	6	2.0	192	2	T05159	hypothetical prote	481	2.0	220	2	S29195	hypothetical prote
409	6	2.0	192	2	A71465	probable translati	482	2.0	220	2	T34306	germin type 2 - Ar
410	6	2.0	193	2	G70884	probable Tsl602res	483	2.0	222	2	S71254	hypothetical prote
411	6	2.0	194	1	S76403	hypothetical prote	484	2.0	222	2	T27842	hypothetical prote
412	6	2.0	194	2	C72554	hypothetical prote	485	2.0	223	2	C72733	hypothetical prote
413	6	2.0	195	2	JC2258	substrate protein	486	2.0	224	2	S51579	cystathionine gamm
414	6	2.0	195	2	B82040	conserved hypothet	487	2.0	225	2	S05546	dehydrin 18 - barl
415	6	2.0	195	2	E83298	hypothetical prote	488	2.0	225	2	E69896	hypothetical prote
416	6	2.0	196	2	H71540	hypothetical prote	489	2.0	226	1	TMVSS	PDGF-related trans
417	6	2.0	196	3	T39648	conserved hypothet	490	2.0	227	2	C75582	conserved hypothet
418	6	2.0	197	2	B64387	imidazoleglycerol-	491	2.0	228	2	S58986	cytochrome-c oxid
419	6	2.0	197	2	T44998	phenylacrylic acid	492	2.0	228	2	S19132	rab25 protein - ri
420	6	2.0	197	2	A70033	carbonic anhydrase	493	2.0	228	2	G64908	probable transcrip
421	6	2.0	197	2	B69447	hypothetical prote	494	2.0	228	2	T22259	hypothetical prote
422	6	2.0	197	2	T10413	hypothetical prote	495	2.0	228	2	T26532	hypothetical prote
423	6	2.0	197	2	T36584	hypothetical prote	496	2.0	228	2	B71283	conserved hypothet
424	6	2.0	198	2	G81675	hypothetical prote	497	2.0	228	2	A72740	hypothetical prote
425	6	2.0	198	2	S07130	riboflavin synthas	498	2.0	229	2	S66342	ribonuclease II (E
426	6	2.0	198	2	B71072	casein B precursor	499	2.0	229	2	E70471	conserved hypothet
427	6	2.0	198	2	A49243	hypothetical prote	500	2.0	230	2	T27498	hypothetical prote
428	6	2.0	199	2	H71519	GLORF-C4 protein -	501	2.0	231	2	S57271	14-3-3 protein hom
429	6	2.0	199	2	S34756	probable riboflavi	502	2.0	231	2	B49934	phosphoglycolate p
430	6	2.0	199	2	T01596	14-3-3 protein eta	503	2.0	231	2	D49934	phosphoglycolate p
431	6	2.0	199	2	C64922	GTP-binding protei	504	2.0	231	2	S17367	rab15B protein - w
432	6	2.0	199	2	T14468	conserved hypothet	505	2.0	231	2	D81441	probable integral
433	6	2.0	199	2	T14468	pollen coat protei	506	2.0	231	2	D65207	hypothetical 26.3K
434	6	2.0	200	2	S64905	hypothetical prote	507	2.0	231	2	S62530	hypothetical prote
435	6	2.0	200	2	G70921	hypothetical prote	508	2.0	232	2	D71442	hypothetical prote
436	6	2.0	201	2	F65137	hypothetical 21.7	509	2.0	232	2	JQ1199	replication protei
437	6	2.0	202	2	H70909	probable resolvase	510	2.0	232	2	T16873	hypothetical prote
438	6	2.0	202	2	T11744	dehydrin - kidney	511	2.0	233	2	T00966	hypothetical prote
439	6	2.0	203	2	T44695	btuR protein (impo	512	2.0	233	2	S15954	repb protein - Lac
440	6	2.0	203	2	H70522	probable resolvase	513	2.0	234	2	E72456	hypothetical prote
441	6	2.0	203	2	JC4645	cardiotrophin-1 -	514	2.0	235	2	A56757	14-3-3 regulatory
442	6	2.0	203	2	S40997	hypothetical prote	515	2.0	235	2	T47432	hypothetical prote
443	6	2.0	203	2	T36102	hypothetical prote	516	2.0	236	2	T16164	hypothetical prote
444	6	2.0	203	2	B70986	probable isomerase	517	2.0	236	2	F71723	ribosomal protein
445	6	2.0	203	2	S60449	probable membrane	518	2.0	238	2	S77699	inner cell wall ma
446	6	2.0	203	2	T12830	hypothetical prote	519	2.0	238	2	F81151	hypothetical prote
447	6	2.0	204	1	FQHUGL	granulocyte colony	520	2.0	238	2	S12259	hypothetical prote
448	6	2.0	205	2	T05713	dehydrin - barley	521	2.0	238	2	T32889	hypothetical prote
449	6	2.0	205	2	T46077	hypothetical prote	522	2.0	239	2	G02630	hypothetical prote
450	6	2.0	205	2	T16560	hypothetical prote	523	2.0	239	2	E75351	conserved hypothet
451	6	2.0	206	1	KASHS1	alpha-si-casein pr	524	2.0	239	2	T21050	3-oxoacyl-lacyl-ca
452	6	2.0	206	2	A45661	alpha-si-casein -	525	2.0	240	2	F70443	hypothetical prote
453	6	2.0	206	2	S52228	ferripyochelin-bin	526	2.0	241	2	H72512	hypothetical prote
454	6	2.0	206	2	D82173	conserved hypothet	527	2.0	242	2	T72512	hypothetical prote
455	6	2.0	206	2	T34347	hypothetical prote	528	2.0	242	2	E83077	conserved hypothet
456	6	2.0	207	2	A24573	granulocyte colony	529	2.0	244	2	T12051	hypothetical prote
457	6	2.0	208	2	H83283	precocurin isomeras	530	2.0	244	2	T19636	hypothetical prote
458	6	2.0	210	2	T49785	hypothetical prote	531	2.0	244	2	A41021	NAD+-asparagine A
459	6	2.0	211	2	S60328	XP group C protein	532	2.0	244	2	S72896	hypothetical prote
460	6	2.0	212	2	S23303	protein kinase C i	533	2.0	245	1	PSHUAM	14-3-3 protein zet
461	6	2.0	212	2	F62581	ribonuclease III X	534	2.0	245	2	A47389	14-3-3 protein zet
462	6	2.0	212	2	A69410	conserved hypothet	535	2.0	245	2	S15076	protein kinase reg
463	6	2.0	213	2	S35522	heterochromatin pr	536	2.0	245	2	S23304	protein kinase C i
464	6	2.0	213	2	T46078	hypothetical prote	537	2.0	245	2	S13467	14-3-3 protein - b
465	6	2.0	213	2	E75525	transcription regu	538	2.0	245	2	JC2502	mitochondrial impo
466	6	2.0	214	2	S34262	alpha-si-casein -	539	2.0	245	2	S59927	14-3-3 protein, th
467	6	2.0	214	2	S06102	alpha-si-casein -	540	2.0	245	2	S65013	14-3-3 protein zet

541	6	2.0	245	2	JC5232	14-3-3 protein iso	614	257	2	C70715	hypothetical prote
542	6	2.0	245	2	JC5384	14-3-3 zeta protei	615	257	2	T32179	hypothetical prote
543	6	2.0	245	2	C65206	thiF protein - Esc	616	258	1	WZBE35	gene 35 protein -
544	6	2.0	246	2	S34755	14-3-3 protein (cl	617	258	2	T07389	14-3-3 protein tft
545	6	2.0	246	2	A49023	nerve growth facto	618	258	2	T04131	14-3-3 protein, is
546	6	2.0	246	2	S38509	14-3-3 protein eta	619	258	2	T08843	14-3-3 protein hom
547	6	2.0	246	2	A60031	14-3-3 protein eta	620	258	2	B61228	collagen alpha 1(I
548	6	2.0	246	2	A40484	14-3-3 protein eta	621	258	2	S71561	drought-induced pr
549	6	2.0	246	2	SL3610	14-3-3 protein - b	622	258	2	T51689	probable transcrip
550	6	2.0	246	2	S38532	protein 14-3-3 eta	623	258	2	T34422	hypothetical prote
551	6	2.0	246	2	T35086	hypothetical prote	624	259	1	A47237	14-3-3 protein hom
552	6	2.0	246	2	T08800	hypothetical prote	625	259	2	T12951	14-3-3-like protei
553	6	2.0	247	1	Q0BPF1	HXIF3 protein prec	626	259	2	S57283	14-3-3 brain prote
554	6	2.0	247	2	A48982	glutathione transf	627	259	2	S64423	probable membrane
555	6	2.0	247	2	B49023	14-3-3 protein gam	628	260	1	S30927	14-3-3 protein hom
556	6	2.0	247	2	T08842	14-3-3 protein hom	629	260	2	T07388	14-3-3 protein tft
557	6	2.0	247	2	G70677	probable echa20 pr	630	260	2	T02051	14-3-3 protein hom
558	6	2.0	248	1	JC1122	14-3-3 protein - f	631	260	2	S57272	14-3-3 protein hom
559	6	2.0	248	1	JQ1680	14-3-3 protein hom	632	260	2	S20580	14-3-3 protein hom
560	6	2.0	248	2	S34753	stratificin - human	633	260	2	T41687	conserved hypothet
561	6	2.0	248	2	S38956	epithelial cell ma	634	261	1	SL8911	14-3-3 protein hom
562	6	2.0	248	2	JC2581	14-3-3 protein - C	635	261	2	S52899	14-3-3 protein hom
563	6	2.0	248	2	S53727	14-3-3 protein hom	636	261	2	S52900	14-3-3 protein hom
564	6	2.0	248	2	T22500	hypothetical prote	637	261	2	T01752	GF14-6 protein - m
565	6	2.0	248	2	T23759	hypothetical prote	638	261	2	D69606	spore coat-associ
566	6	2.0	248	2	T14957	hypothetical prote	639	261	2	T10293	hypothetical prote
567	6	2.0	248	2	T02554	hypothetical prote	640	261	2	JQ2028	matrix-associated
568	6	2.0	249	2	T07383	14-3-3 protein tft	641	262	1	NDKE2M	type II site-speci
569	6	2.0	249	2	T04128	14-3-3 protein, is	642	262	2	A36024	sepiapterin reduct
570	6	2.0	249	2	F71311	probable triosepho	643	262	2	T04152	GF14-b protein - r
571	6	2.0	249	2	T36222	hypothetical prote	644	262	2	T04406	14-3-3b protein -
572	6	2.0	249	2	G75583	hypothetical prote	645	262	2	S57276	14-3-3 protein hom
573	6	2.0	249	2	T04939	hypothetical prote	646	262	2	S72890	hypothetical prote
574	6	2.0	249	2	T46184	hypothetical prote	647	263	2	T06203	14-3-3 protein - b
575	6	2.0	249	2	T21083	hypothetical prote	648	263	2	T22135	hypothetical prote
576	6	2.0	250	2	T36428	hypothetical prote	649	264	2	T12572	14-3-3 protein - c
577	6	2.0	250	2	T10663	hypothetical prote	650	264	2	A31638	transformer-2 sex-
578	6	2.0	250	2	S64849	hypothetical prote	651	264	2	T51337	auxin-induced prot
579	6	2.0	251	2	T04101	14-3-3 protein ho	652	264	2	T23866	hypothetical prote
580	6	2.0	251	2	S47370	14-3-3 protein hom	653	265	1	MNVURV	nonstructural prot
581	6	2.0	251	2	H70666	hypothetical prote	654	265	2	T04154	GF14-d protein - r
582	6	2.0	251	2	B70765	probable coBM - My	655	265	2	T36391	hypothetical prote
583	6	2.0	251	2	A38912	NAD+-asparagine A	656	265	2	T33523	hypothetical prote
584	6	2.0	252	2	H83062	molybdopterin bios	657	265	2	S39874	hypothetical prote
585	6	2.0	252	2	S30292	mating-type switch	658	266	2	H83517	probable transcrip
586	6	2.0	253	2	T27951	hypothetical prote	659	267	1	S30863	BMH1 protein - yea
587	6	2.0	253	2	T04967	hypothetical prote	660	267	1	CDPU2R	chlorophyll a/b-bi
588	6	2.0	253	2	F72388	prephenate dehydro	661	268	2	S71173	14-3-3 protein hom
589	6	2.0	254	2	T07103	14-3-3 protein hom	662	268	2	T51388	14-3-3-LIKE PROTEI
590	6	2.0	254	2	T07387	14-3-3 protein tft	663	268	2	B42424	chitinase (PC 3.2
591	6	2.0	254	2	S55375	14-3-3 protein - p	664	268	2	B75279	survival protein S
592	6	2.0	255	1	S31375	14-3-3 protein eps	665	268	2	F65134	hypothetical 30.0K
593	6	2.0	255	1	QRECH3	membrane-bound iro	666	268	2	T02448	hypothetical prote
594	6	2.0	255	2	T04127	14-3-3 protein, is	667	268	2	S24377	hypothetical prote
595	6	2.0	255	2	T02050	14-3-3 protein hom	668	269	1	CDPM80	hypothetical prote
596	6	2.0	255	2	JX0341	14-3-3 protein eps	669	269	2	H75576	chlorophyll a/b-bi
597	6	2.0	255	2	S57277	14-3-3 protein hom	670	270	2	T39156	cobalamin synthase
598	6	2.0	255	2	S47969	14-3-3 protein hom	671	270	2	T37814	dna damage checkpo
599	6	2.0	255	2	I38947	14-3-3 protein eps	672	270	2	T43316	DNA damage checkpo
600	6	2.0	255	2	B71490	hypothetical prote	673	271	2	T43039	hypothetical prote
601	6	2.0	256	2	T04153	GF14-c protein - r	674	271	2	T45211	DNA damage checkpo
602	6	2.0	256	2	JC7180	14-3-3 protein hom	675	271	2	A25669	PDGF-related trans
603	6	2.0	256	2	A32373	transformer-2 prot	676	271	2	T36885	probable membrane
604	6	2.0	256	2	H71537	hypothetical prote	677	272	2	T04129	14-3-3 protein, is
605	6	2.0	256	2	H82517	hypothetical prote	678	272	2	D75626	glycerol-3-phospha
606	6	2.0	256	2	B82661	hypothetical prote	679	272	2	C75291	guanylate kinase -
607	6	2.0	256	2	B82648	hypothetical prote	680	273	2	S51250	BMH2 protein - yea
608	6	2.0	256	2	T16661	hypothetical prote	681	273	2	T49994	14-3-3-like protei
609	6	2.0	257	1	S40985	H+-transporting AT	682	273	2	H70849	hypothetical prote
610	6	2.0	257	2	T13335	alcohol dehydrogen	683	273	2	S28132	gas vesicle protei
611	6	2.0	257	2	T08840	14-3-3 protein hom	684	275	1	S38861	14-3-3 protein hom
612	6	2.0	257	2	G64181	ABC-type transport	685	275	2	D72854	nuclear matrix ass
613	6	2.0	257	2	T33648	hypothetical prote	686	275	2	D65100	hypothetical L-ser

687	6	2.0	276	1	JC4005	glutamate racemase	760	2.0	306	2	T06607	hypothetical prote
688	6	2.0	276	2	F83296	conserved hypothet	761	2.0	307	2	T43073	ts6770 transposase
689	6	2.0	276	2	H75588	conserved hypothet	762	2.0	307	2	H82743	methionyl-tRNA for
690	6	2.0	276	2	T46368	hypothetical prote	763	2.0	308	2	T44427	hypothetical prote
691	6	2.0	278	2	T36647	probable hydrolase	764	2.0	308	2	T43293	multidrug resistan
692	6	2.0	278	2	C83280	probable transcrip	765	2.0	308	2	S71746	cwl1 protein - fis
693	6	2.0	279	2	T16201	hypothetical prote	766	2.0	309	2	T41203	WD repeat protein
694	6	2.0	280	2	F75057	hypothetical prote	767	2.0	310	2	T34809	ribosomal protein
695	6	2.0	281	2	C72279	sugar ABC transpor	768	2.0	310	2	T33611	N-acetyl-gamma-glu
696	6	2.0	282	2	T21222	hypothetical prote	769	2.0	311	2	T43947	hypothetical prote
697	6	2.0	283	2	A60364	tropomyosin - mlgr	770	2.0	311	2	T23846	hypothetical prote
698	6	2.0	283	2	D71130	probable oligopept	771	2.0	312	1	CRH04	carbonate dehydrat
699	6	2.0	284	1	KART	alpha-casein precu	772	2.0	312	2	G71131	probable ATP-bindi
700	6	2.0	284	2	T06129	zinc finger protei	773	2.0	312	2	E83624	probable transcrip
701	6	2.0	285	1	A37757	acetylactate decar	774	2.0	313	2	T35826	probable dehydroge
702	6	2.0	285	2	E83234	hypothetical prote	775	2.0	314	2	E64916	hypothetical prote
703	6	2.0	286	2	S73424	spermidine/putresc	776	2.0	314	2	E83487	hypothetical prote
704	6	2.0	286	2	A54070	3-hydroxyanthranil	777	2.0	315	2	C75626	1-phosphoethanol
705	6	2.0	286	2	C61615	sericin MG-2 - gre	778	2.0	315	2	T06053	1-phosphoethanol
706	6	2.0	287	2	E70925	probable ribosomal	779	2.0	315	2	T51683	myb-related trans
707	6	2.0	287	2	JH0332	IgA (Fc) receptor,	780	2.0	315	2	S53589	SOL2 protein - yea
708	6	2.0	288	2	S41382	pobR protein - Pse	781	2.0	316	1	A39407	1-phosphoethanol
709	6	2.0	288	2	S57706	hypothetical prote	782	2.0	316	1	T34838	probable transfe
710	6	2.0	288	2	F83613	probable transcrip	783	2.0	316	2	T42012	hypothetical prote
711	6	2.0	289	2	B82534	heat shock protein	784	2.0	317	2	B82637	bi-functional bioti
712	6	2.0	289	2	S35022	nolt protein - Rhi	785	2.0	317	2	T17792	hypothetical prote
713	6	2.0	290	2	T21868	hypothetical prote	786	2.0	317	2	D72700	hypothetical prote
714	6	2.0	291	2	S22041	hypothetical prote	787	2.0	317	2	I37371	ER calcium-binding
715	6	2.0	291	2	S48252	serine-rich protei	788	2.0	317	2	I52575	gene MRK protein -
716	6	2.0	291	2	T04334	DNA binding protei	789	2.0	317	2	T39869	hypothetical prote
717	6	2.0	292	2	T04637	40s ribosomal prot	790	2.0	318	1	QXBO1M	NADH dehydrogen
718	6	2.0	292	2	I51171	transcription fact	791	2.0	318	2	S48455	probable membrane
719	6	2.0	292	2	T07021	hypothetical prote	792	2.0	319	2	S72379	transposase - Ente
720	6	2.0	292	2	T08862	hypothetical prote	793	2.0	319	2	C64852	probable pseudouri
721	6	2.0	292	2	T32178	hypothetical prote	794	2.0	319	2	UC5402	vitamin D receptor
722	6	2.0	293	2	S23384	cyclohexadienyl de	795	2.0	319	2	T05584	hypothetical prote
723	6	2.0	294	2	T35492	probable endonucle	796	2.0	320	2	S42651	hypothetical prote
724	6	2.0	294	2	T29028	hypothetical prote	797	2.0	320	2	A75278	hypothetical prote
725	6	2.0	295	2	T35330	probable RNA polym	798	2.0	320	2	C82170	peptide ABC transp
726	6	2.0	295	2	S67860	gumK protein - Xan	799	2.0	320	2	T06729	hypothetical prote
727	6	2.0	295	2	T00725	hypothetical prote	800	2.0	321	2	T02520	hypothetical prote
728	6	2.0	295	2	T04483	probable ring fing	801	2.0	321	2	T02987	myb-related protei
729	6	2.0	295	2	E82659	peptidyl-prolyl ci	802	2.0	321	2	F83358	hypothetical prote
730	6	2.0	296	2	F82077	nicotinate-nucleot	803	2.0	322	1	ATBP1	antipressor prot
731	6	2.0	296	2	B70745	hypothetical prote	804	2.0	322	2	S15201	hydrogenase expres
732	6	2.0	296	2	S57179	hypothetical prote	805	2.0	322	2	T43601	yop targeted effec
733	6	2.0	298	2	B64674	conserved hypothet	806	2.0	323	2	S20099	transforming prote
734	6	2.0	298	2	G71841	hypothetical prote	807	2.0	323	2	F71130	probable oligopept
735	6	2.0	298	2	T25695	hypothetical prote	808	2.0	323	2	D70125	Lambda CII stabili
736	6	2.0	300	2	T25275	hypothetical prote	809	2.0	323	2	T08715	hypothetical prote
737	6	2.0	300	2	B70725	probable linB prot	810	2.0	324	2	S18547	alkanal monooxygen
738	6	2.0	300	2	F75586	urea/short-chain a	811	2.0	325	2	S46760	hypothetical prote
739	6	2.0	300	2	T29112	hypothetical prote	812	2.0	325	2	T24689	hypothetical prote
740	6	2.0	301	2	T17321	hypothetical prote	813	2.0	327	2	A46484	apoptosis-mediati
741	6	2.0	301	2	A58928	phosducin-like pro	814	2.0	327	1	IMBKB	immunogenic protei
742	6	2.0	301	2	C72616	hypothetical prote	815	2.0	329	2	T35567	probable pantheon
743	6	2.0	301	2	G83344	probable permease	816	2.0	329	2	T36392	nicotinate-nucleot
744	6	2.0	301	2	T33702	hypothetical prote	817	2.0	329	2	F75494	hypothetical prote
745	6	2.0	302	2	T01540	hypothetical prote	818	2.0	329	2	T44459	arginine metabolis
746	6	2.0	303	2	S29567	cysteine synthase	819	2.0	329	2	B83535	transcription regu
747	6	2.0	303	2	T02588	hypothetical prote	820	2.0	329	2	T13016	hypothetical prote
748	6	2.0	304	2	E70698	hypothetical prote	821	2.0	330	2	T25169	farnesyltransf
749	6	2.0	304	2	A83309	hypothetical prote	822	2.0	330	2	A54058	high-affinity nick
750	6	2.0	304	2	H82964	oxidoreductase Rmd	823	2.0	331	2	F71943	nickel transport p
751	6	2.0	305	2	A56554	transcription fact	824	2.0	331	2	S53113	conserved hypothet
752	6	2.0	305	2	H67579	conserved hypothet	825	2.0	332	2	F75473	probable integrat
753	6	2.0	305	2	T34867	probable chitosana	826	2.0	334	2	T29061	COP1-interacting p
754	6	2.0	306	1	QREBOB	oligopeptide trans	827	2.0	334	2	T51245	hypothetical prote
755	6	2.0	306	2	B36263	oligopeptide trans	828	2.0	334	2	T02674	hypothetical prote
756	6	2.0	306	2	E64184	oligopeptide trans	829	2.0	335	2	T41037	hypothetical prote
757	6	2.0	306	2	S33538	embryogenic cell p	830	2.0	335	2	S08341	myristylated alani
758	6	2.0	306	2	B71439	hypothetical prote	831	2.0	335	2	A41845	orf A - Treponema
759	6	2.0	306	2	C82242	oligopeptide ABC t	832	2.0	335	2		

conserved hypothet
plasmaenogen activa
hypothetical prote
hypothetical prote
beta-ketoadipyl Co
plasmaenogen activa
probable carboxyl-
probable MFS trans
hypothetical prote
tyrosine-specific
ATP-dependent Clp
hypothetical prote
hypothetical prote
tylosin biosynthes
translation elonga
hypothetical prote
hypothetical prote
protein C inhibito
hypothetical prote
type II secretion
hypothetical 44.9K
outf protein - Erw

ALIGNMENTS

RESULT 1
H83483
probable transmembrane sensor PA1301 [Imported] - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 15-Sep-2000
C:Accession: H83483
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A:Reference number: A82950
A:Accession: H83483
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-327 <STO>
A:Cross-references: GB:AE004559; GB:AE004091; NID:g9947228; PIDN:AAG04690.1; GSPDB:GN001
A:Experimental source: strain PAO1
C:Genetics:
A:Gene: PA1301

Query Match 3.0%; Score 9; DB 2; Length 327;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 16 LALPALLPV 24
|||||||
Db 304 LALPALLPV 312

RESULT 2
KGRT
gamma-casein precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 13-Jun-1983 #sequence_revision 13-Jun-1983 #text_change 31-May-1996
C:Accession: A03111
R:Hobbs, A.A.; Rosen, J.M.
Nucleic Acids Res. 10, 8079-8098, 1982
A:Title: Sequence of rat alpha- and gamma-casein mRNAs: evolutionary comparison of the c
A:Reference number: A93452; MUID:83143278
A:Accession: A03111
A:Molecule type: mRNA
A:Residues: 1-179 <HOB>
C:Superfamily: gamma-casein
C:Keywords: phosphoprotein

F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-179/Product: gamma-casein #status predicted <MAT>

Query Match 2.7%; Score 8; DB 1; Length 179;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSSE 166

|||||||

Db 50 SASSSSSE 57

RESULT 3

T49104

hypothetical protein AT4g21970 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000

C:Accession: T49104

R:Bevan, M.; Wedler, H.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Lemcke, K.

submitted to the Protein Sequence Database, May 2000

A:Reference number: Z25016

A:Accession: T49104

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-181 <BEV>

A:Cross-references: EMBL:AL022140; GSPDB:GN00062; ATSP:AT4g21970

A:Experimental source: cultivar Columbia; BAC clone F1N20

C:Genetics:

A:Gene: ATSP:AT4g21970

A:Map position: 4

A:Introns: 142/1

Query Match 2.7%; Score 8; DB 2; Length 181;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 158 FSASSSSS 165

|||||||

Db 35 FSASSSSS 42

RESULT 4

I49685

gamma-casein precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 13-Aug-1999

C:Accession: I49685

R:Sasaki, T.; Sasaki, M.; Enami, J.

Zool. Sci. 10, 65-72, 1993

A:Title: Mouse gamma-casein cDNA: PCR cloning and sequence analysis.

A:Reference number: I49685; MUID:93320737

A:Accession: I49685

A:Status: preliminary; translated from GB/EMBL/DBBJ

A:Molecule type: mRNA

A:Residues: 1-184 <RES>

A:Cross-references: GB:D10215; NID:g220404; PIDN:BAA01067.1; PID:g220405

C:Superfamily: gamma-casein

Query Match 2.7%; Score 8; DB 2; Length 184;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSSE 166

|||||||

Db 51 SASSSSSE 58

RESULT 5

T05479

hypothetical protein T805.180 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
 C:Accession: T05479
 R:Bevan, M.; Wedler, H.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Schueller
 submitted to the Protein Sequence Database, February 1998
 A:Reference number: Z15417
 A:Accession: T05479
 A:Molecule type: DNA
 A:Residues: 1-241 <BEV>
 A:Cross-references: EMBL:AL021890
 A:Experimental source: cultivar Columbia; BAC clone T805
 C:Genetics:
 A:Map position: 4
 A:Introns: 142/1; 169/3; 193/1; 211/1; 223/3
 A:Note: T805.180

Query Match 2.7%; Score 8; DB 2; Length 241;
 Best Local Similarity 100.0%; Pred. No. 7.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 158 FSASSSS 165
 |||||
 Db 35 FSASSSS 42

RESULT 6
 T27654
 hypothetical protein ZK1025.9 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T27654
 R:Lennard, N.
 submitted to the EMBL Data Library, March 1998
 A:Reference number: Z20400
 A:Accession: T27654
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-341 <WIL>
 A:Cross-references: EMBL:AL022288; PIDN:CAA18368.1; GSPDB:GN00019; CESP:ZK1025.9
 A:Experimental source: clone ZK1025
 C:Genetics:
 A:Gene: CESP:ZK1025.9
 A:Map position: 1
 A:Introns: 6/1; 54/1; 76/1; 97/1; 158/3; 232/2; 324/3

Query Match 2.7%; Score 8; DB 2; Length 341;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 159 SASSSSE 166
 |||||
 Db 103 SASSSSE 110

RESULT 7
 T50912
 hypothetical protein ORF440 [imported] - Rubrivivax gelatinosus
 C:Species: Rubrivivax gelatinosus
 C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
 C:Accession: T50912
 R:Nagashima, K.V.; Igarashi, N.; Harada, J.; Nagashima, S.; Matsuura, K.; Shimada, K.
 submitted to the EMBL Data Library, November 1999
 A:Description: Determination of Nucleotide Sequences of Rubrivivax gelatinosus Photosynt
 A:Reference number: Z25270
 A:Accession: T50912
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-440 <NAG>
 A:Cross-references: EMBL:AB034704; PIDN:BAA94065.1
 A:Experimental source: strain IL144
 C:Genetics:

A:Note: ORF440

Query Match 2.7%; Score 8; DB 2; Length 440;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LALPALLP 23
 |||||
 Db 42 LALPALLP 49

RESULT 8
 S60028
 ferredoxin--NADP+ reductase (EC 1.18.1.2) precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: S60028; I49671
 R:Itoh, S.; Iemura, O.; Yamada, E.; Yoshimura, T.; Tsujikawa, K.; Kohama, Y.; Mimura,
 Biochim. Biophys. Acta 1264, 159-162, 1995
 A:Title: cDNA cloning of mouse ferredoxin reductase from kidney.
 A:Reference number: I49671; MUID:96085117
 A:Accession: S60028
 A:Molecule type: mRNA
 A:Residues: 1-494 <ITO>
 A:Cross-references: EMBL:D49920; NID:g1088468; PIDN:BAA08659.1; PID:g1088469
 C:Genetics:
 A:Genome: nuclear
 C:Superfamily: human ferredoxin--NADP+ reductase
 C:Keywords: FAD; mitochondrion; NADP; oxidoreductase
 F:1-34/Domain: transit peptide (mitochondrion) #status predicted <TNP>
 F:35-494/Product: ferredoxin--NADP+ reductase #status predicted <MAT>

Query Match 2.7%; Score 8; DB 1; Length 494;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 258 RRLTELL 265
 |||||
 Db 276 RRLTELL 283

RESULT 9
 B64905
 sugar kinase homolog ydev - Escherichia coli
 C:Species: Escherichia coli
 C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 17-Mar-2000
 C:Accession: B64905
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
 A.: Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617
 A:Accession: B64905
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-530 <BLAT>
 A:Cross-references: GB:AE000248; GB:U00096; NID:g1787783; PIDN:AAC74584.1; PID:g17877
 A:Experimental source: strain K-12, substrain MG1655
 C:Genetics:
 A:Gene: ydev
 C:Superfamily: xylulokinase

Query Match 2.7%; Score 8; DB 2; Length 530;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 AETGERLV 48
 |||||
 Db 475 AETGERLV 482

```
RESULT 10
S54098
1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase (EC 3.1.4.11) - Arabidopsis th
N:Alternate names: protein T21L8.40
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 10-Sep-1999 #sequence_revision 26-May-2000 #text_change 26-May-2000
C:Accession: T12984; S54098
R:Cholsne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; Sa
submitted to the Protein Sequence Database, July 1999
A:Reference number: Z17586
A:Accession: T12984
A:Molecule type: DNA
A:Residues: 1-531 <CHO>
A:Cross-references: EMBL:AL096860; GSPDB:GN00061; ATSP:T21L8.40
A:Experimental source: cultivar Columbia; BAC clone T21L8
R:Kopka, J.; Willmitzer, L.; Mueller-Roeber, B.
submitted to the EMBL Data Library, March 1995
A:Reference number: S54098
A:Accession: S54098
A:Molecule type: mRNA
A:Residues: 1-81, K', 83-357, D', 359-531 <KOP>
A:Cross-references: EMBL:X85973; NID:g769742; PIDN:CAAS9962.1; PID:g769743
C:Genetics:
A:Gene: Plc2; ATSP:T21L8.40
A:Map position: 3
A:Introns: 102/3; 166/2; 211/3; 317/2; 356/3; 406/3; 434/3
C:Superfamily: Arabidopsis thaliana 1-phosphatidylinositol-4,5-bisphosphate phosphodiester
bisphosphate phosphodiesterase domain Y homology
C:Keywords: phosphoric diester hydrolase
F:107-247/Domain: 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase domain X hom
Query Match 2.7%; Score 8; DB 1; Length 531;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 259 RRLTELLG 266
Db 38 RRLTELLG 45
RESULT 11
JC5119
anti-mullerian hormone homolog precursor - chicken
C:Species: Gallus gallus (chicken)
C:Date: 02-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 04-Mar-2000
R:Nepper, M.; Lowe, R.; Galuska, S.; Hofmann, K.J.; Smith, R.G.; Elbrecht, A.
Gene 176, 203-209, 1996
A:Title: Molecular cloning of an avian anti-Muellerian hormone homologue.
A:Reference number: JC5119; MUID:97075931
A:Accession: JC5119
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-644 <NEE>
A:Cross-references: GB:U61754; NID:g1432157; PID:g1432158
C:Comment: This protein is a glycoprotein belonging to the transforming growth factor be
ng embryonic development and sex determination.
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-644/Product: anti-Muellerian hormone homologue #status predicted <MAT>
Query Match 2.7%; Score 8; DB 2; Length 644;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 200 PLSTRVPG 207
Db 579 PLSTRVPG 586
RESULT 12
```

```
T49472
hormone-sensitive lipase related protein [imported] - Neurospora crassa
N:Alternate names: protein B14D6.240
C:Species: Neurospora crassa
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C:Accession: T49472
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatu
submitted to the Protein Sequence Database, May 2000
A:Reference number: 225022
A:Accession: T49472
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-781 <SCH>
A:Cross-references: EMBL:AL356173; GSPDB:GN00116; NCSP:B14D6.240
A:Experimental source: BAC clone B14D6; strain OR74A
C:Genetics:
A:Gene: NCSP:B14D6.240
A:Map position: 6
Query Match 2.7%; Score 8; DB 2; Length 781;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 269 DGALLVRL 276
Db 749 DGALLVRL 756
RESULT 13
TSHUP2
thrombospondin 2 precursor - human
C:Species: Homo sapiens (man)
C:Date: 19-May-1995 #sequence_revision 03-Aug-1995 #text_change 13-Aug-1999
C:Accession: A47379; A42173
R:LaBell, T.L.; Byers, P.H.
Genomics 17, 225-229, 1993
A:Title: Sequence and characterization of the complete human thrombospondin 2 cDNA: p
A:Reference number: A47379; MUID:94010892
A:Accession: A47379
A:Molecule type: mRNA
A:Residues: 1-1172 <LAB>
A:Cross-references: GB:L12350; NID:g307505; PIDN:AAA03703.1; PID:g307506
R:LaBell, T.L.; Milewicz, D.J.; Disceche, C.M.; Byers, P.H.
Genomics 12, 421-429, 1992
A:Title: Thrombospondin II: partial cDNA sequence, chromosome location, and expressio
A:Reference number: A42173; MUID:92217961
A:Accession: A42173
A:Molecule type: mRNA
A:Residues: 560-1172 <LA2>
A:Cross-references: GB:M81339
A:Experimental source: fibroblast
A:Note: sequence extracted from NCBI backbone (NCBIN:95091, NCBIP:95096)
C:Genetics:
A:Gene: GDB:THBS2; TSP2
A:Cross-references: GDB:128789; OMIM:188061
A:Map position: 6q27-6q27
C:Complex: homotrimer, disulfide linked
C:Function:
A:Description: participates in cell migration and adhesion, and in platelet aggregati
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology;
C:Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; tri
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-1172/Product: thrombospondin 2 #status predicted <MAT>
F:319-377/Domain: von Willebrand factor type C repeat homology <VWC>
F:380-431/Domain: thrombospondin type 1 repeat homology <THR1>
F:436-492/Domain: thrombospondin type 1 repeat homology <THR2>
F:493-549/Domain: thrombospondin type 1 repeat homology <THR3>
F:553-588/Domain: EGF homology <EGF1>
F:652-691/Domain: EGF homology <EGF>
F:928-930/Region: cell attachment (R-G-D) motif
F:151,316,330,457,584,710,1069/Binding site: carbohydrate (Asn) (covalent) #status pr
F:167-226/Disulfide bonds: #status predicted
```

F:266,270/Disulfide bonds: Interchain #status predicted
F:612/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match 2.7%; Score 8; DB 1; Length 1172;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ALEGPGLS 10
|||||
DB 102 ALEGPGLS 109

RESULT 14

secretin - guinea pig
C:Species: Cavia porcellus (guinea pig)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: B57082
R:Bussacil, L.; Cauvin, A.; Gourlet, P.; Gossen, D.; de Neef, P.; Rathe, J.; Robberecht, B.
Biochim. Biophys. Acta 1038, 355-359, 1990
A:Title: Purification and amino acid sequence of vasoactive intestinal peptide, peptide
A:Reference number: S09688; MUID:90254163
A:Accession: B57082
A:Molecule type: protein
A:Residues: 1-26 <BUS>
C:Superfamily: glucagon
C:Keywords: amidated carboxyl end; duodenal mucosa; duplication; hormone; secretagogue
F:1-26/product: secretin #status experimental <MAT>
F:267/Modified site: amidated carboxyl end (Val) #status experimental

Query Match 2.3%; Score 7; DB 1; Length 26;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLQ 234
|||||
DB 17 RLQRLQ 23

RESULT 15

secretin - human
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S07443
R:Carlquist, M.; Joernvall, H.; Forssmann, W.G.; Thulin, L.; Johansson, C.; Mutt, V.
IRCS Med. Sci. 13, 217-218, 1985
A:Title: Human secretin is not identical to the porcine/bovine hormone.
A:Reference number: S07443
A:Accession: S07443
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-27 <CAR>
C:Genetics:
A:Gene: GDB:SCT
A:Cross-references: GDB:270550
A:Map position: Xp21.1-Xp21.1
C:Superfamily: glucagon
C:Keywords: amidated carboxyl end; duplication
F:27/Modified site: amidated carboxyl end (Val) #status predicted

Query Match 2.3%; Score 7; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLQ 234
|||||
DB 18 RLQRLQ 24

RESULT 16

secretin - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 20-Mar-1998
C:Accession: A91291; A01544
R:Carlquist, M.; Joernvall, H.; Mutt, V.
FEBS Lett. 127, 71-74, 1981
A:Title: Isolation and amino acid sequence of bovine secretin.
A:Reference number: A91291; MUID:81237102
A:Accession: A91291
A:Molecule type: protein
A:Residues: 1-27 <CAR>
C:Superfamily: glucagon
C:Keywords: amidated carboxyl end; duodenal mucosa; duplication; hormone; secretagogue
F:27/Modified site: amidated carboxyl end (Val) #status experimental

Query Match 2.3%; Score 7; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLQ 234
|||||
DB 18 RLQRLQ 24

RESULT 17

secretin - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 20-Mar-1998
C:Accession: C60072
R:Boujoua, Y.; Vandermeers, A.; Robberecht, P.; Vandermeers-Piret, M.C.; Christophe, Regul. Pept. 32, 169-179, 1991
A:Title: Purification and amino acid sequence of vasoactive intestinal peptide, pepti
A:Reference number: A60072; MUID:91239834
A:Accession: C60072
A:Molecule type: protein
A:Residues: 1-27 <BOU>
C:Superfamily: glucagon
C:Keywords: amidated carboxyl end; duplication; hormone; intestine
F:27/Modified site: amidated carboxyl end (Val) #status experimental

Query Match 2.3%; Score 7; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLQ 234
|||||
DB 18 RLQRLQ 24

RESULT 18

secretin - dog
C:Species: Canis lupus familiaris (dog)
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 21-Nov-1997
C:Accession: A27267
R:Shinomura, Y.; Eng, J.; Yalow, R.S.
Life Sci. 41, 1243-1248, 1987
A:Title: Dog secretin: sequence and biologic activity.
A:Reference number: A27267; MUID:87314204
A:Accession: A27267
A:Molecule type: protein
A:Residues: 1-27 <SHI>
A:Experimental source: intestine
C:Superfamily: glucagon
C:Keywords: duplication

Query Match 2.3%; Score 7; DB 2; Length 27;

```

Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLQ 234
    |||||
Db 18 RLQRLQ 24

RESULT 19
C60415
secretin - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 03-Feb-1993 #sequence_revision 03-Feb-1993 #text_change 20-Mar-1998
C:Accession: C60415
R:Gossen, D.; Busecail, L.; Cauvin, A.; Gourlet, P.; De Neef, P.; Rathe, J.; Robberecht,
Peptides 11, 123-128, 1990
A:Title: Amino acid sequence of VIP, PHI and secretin from the rabbit small intestine.
A:Reference number: A60415; MUID:90259845
A:Accession: C60415
A:Molecule type: protein
A:Residues: 1-27 <GOS>
C:Superfamily: glucagon
C:Keywords: amidated carboxyl end; duplication; hormone; intestine; secretagogue
F:27/Modified site: amidated carboxyl end (Leu) #status experimental

Query Match 2.3%; Score 7; DB 2; Length 27;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLQ 234
    |||||
Db 18 RLQRLQ 24

RESULT 20
S13515
retinoic acid receptor beta-1 - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Aug-1999
C:Accession: S13515
R:Zelent, A.; Mendelsohn, C.; Kastner, P.; Krust, A.; Garnier, J.M.; Ruffenach, F.; Le
EMBO J. 10, 71-81, 1991
A:Title: Differentially expressed isoforms of the mouse retinoic acid receptor beta are
A:Reference number: S13515; MUID:91114714
A:Accession: S13515
A:Molecule type: mRNA
A:Residues: 1-74 <ZEL>
A:Cross-references: EMBL:X55569; NID:g53901; PIDN:CAA39918.1; PID:g53902
C:Superfamily: retinoic acid receptor alpha; erba transforming protein Homology
C:Keywords: DNA binding; nucleus; transcription regulation; zinc finger

Query Match 2.3%; Score 7; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 PVPVVRG 29
    |||||
Db 9 PVPVVRG 15

RESULT 21
S13517
retinoic acid receptor beta-3 - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Aug-1999
C:Accession: S13517
R:Zelent, A.; Mendelsohn, C.; Kastner, P.; Krust, A.; Garnier, J.M.; Ruffenach, F.; Le
EMBO J. 10, 71-81, 1991
A:Title: Differentially expressed isoforms of the mouse retinoic acid receptor beta are
A:Reference number: S13515; MUID:91114714
A:Accession: S13517

```

A:Molecule type: DNA
A:Residues: 1-112 <KAW>
A:Cross-references: DDBJ:AP000063; NID:g5105654; PIDN:BAAR1007.1; PID:dl044793; PID:g5105654
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1997

Query Match 2.3%; Score 7; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 159 SASSSS 165
| | | | |
Db 3 SASSSS 9

RESULT 24
IMECE1

colicin E1 immunity protein - Escherichia coli plasmid colicin E1
C:Species: Escherichia coli
C:Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 24-Sep-1999
C:Accession: A94079; B93118; I54985; A03513; A24685
R:Waleh, N.S.; Johnson, P.H.
Proc. Natl. Acad. Sci. U.S.A. 82, 8389-8393, 1985
A:Title: Structural and functional organization of the colicin E1 operon.
A:Reference number: A94079; MUID:86094231
A:Accession: A94079

A:Molecule type: DNA
A:Residues: 1-113 <WAL>
R:Oka, A.; Nomura, N.; Morita, M.; Sugisaki, H.; Sugimoto, K.; Takanami, M.
Mol. Gen. Genet. 172, 151-159, 1979
A:Title: Nucleotide sequence of small ColE1 derivatives: structure of the regions essential for replication.
A:Reference number: A93118; MUID:80010893
A:Accession: B93118

A:Molecule type: DNA
A:Residues: 1-113 <OKA>
A:Cross-references: GB:V00268; NID:g41124; PIDN:CAA23529.1; PID:g41125
A:Note: plasmid pAO3 is a small col E1 derivative
R:Jilk, R.A.; Makris, J.C.; Borchardt, L.; Reznikoff, W.S.
J. Bacteriol. 175, 1264-1271, 1993
A:Title: Implications of Tn5-associated adjacent deletions.
A:Reference number: I54985; MUID:93186691
A:Accession: I54985

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-113 <RES>
A:Cross-references: GB:S56312; NID:g298343
C:Genetics:

A:Gene: imm

A:Map position: 75-100/0-1

A:Genome: plasmid

C:Superfamily: colicin E1 immunity protein

Query Match 2.3%; Score 7; DB 1; Length 113;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 LLCLVLA 17
| | | | |
Db 92 LLCLVLA 98

RESULT 25
I64785

imm protein - Escherichia coli plasmid R100
C:Species: Escherichia coli
C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 26-Aug-1999
C:Accession: I64785
R:Ohtsubo, H.; Ryder, T.B.; Maeda, Y.; Armstrong, K.A.; Ohtsubo, E.
Adv. Biophys. 21, 115-133, 1986
A:Title: DNA replication of the resistance plasmid R100 and its control.

A:Reference number: I51821; MUID:86319522

A:Accession: I64785

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-113 <RES>

A:Cross-references: GB:M26840; NID:g151770; PIDN:AAA26071.1; PID:g151777

C:Genetics:

A:Genome: plasmid

C:Superfamily: colicin E1 immunity protein

Query Match 2.3%; Score 7; DB 2; Length 113;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 LLCLVLA 17
| | | | |
Db 92 LLCLVLA 98

RESULT 26

colicin E1 immunity protein - Shigella sonnei plasmid pKY-1

C:Species: Shigella sonnei

C:Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 26-Aug-1999

C:Accession: S11532

R:Higashi, M.; Hata, M.; Hase, T.; Yamaguchi, K.; Masamune, Y.

J. Gen. Appl. Microbiol. 32, 433-442, 1986

A:Title: The nucleotide sequence of cea and the region of origin of plasmid pKY-1.
A:Reference number: S06218
A:Molecule type: DNA

A:Residues: 1-113 <UCE>

A:Cross-references: EMBL:M37218; NID:g149346; PIDN:AAA98157.1; PID:g149348

C:Genetics:

A:Gene: imm

A:Genome: plasmid

C:Superfamily: colicin E1 immunity protein

Query Match 2.3%; Score 7; DB 2; Length 113;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 LLCLVLA 17
| | | | |
Db 93 LLCLVLA 99

RESULT 27

S44660

ZK353.5 protein - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 25-Dec-1994 #sequence_revision 12-May-1995 #text_change 20-Sep-1999

C:Accession: S44660

R:Favella, A.D.

submitted to the EMBL Data Library, May 1993

A:Description: Sequence of the C. elegans cosmid ZK353.

A:Reference number: S44617

A:Accession: S44660

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-114 <FAV>

A:Cross-references: EMBL:L15313; NID:g289756; PID:g289763

C:Genetics:

A:Introns: 44/1

C:Superfamily: Caenorhabditis elegans ZK353.5 protein

Query Match 2.3%; Score 7; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
|||||||
Db 11 SASSSSS 17

RESULT 28

S27476
hypothetical protein 1 - Azospirillum brasilense
C:Species: Azospirillum brasilense
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 08-Oct-1999
C:Accession: S27476
R:Passaglia, L.M.P.; Nunes, C.P.; Zaha, A.; Schrank, I.S.
submitted to the EMBL Data Library, April 1991
A:Description: The nifHDK operon in the free-living nitrogen fixing bacteria Azospirillum
A:Reference number: S27474
A:Accession: S27476
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-118 <PAS>
A:Cross-references: EMBL:M64344; NID:g142415; PIDN:AAB02345.1; PID:g142419

Query Match 2.3%; Score 7; DB 2; Length 118;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
|||||||
Db 73 SASSSSS 79

RESULT 29

SEPC
secretin precursor - pig
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 24-Apr-1984 #sequence_revision 12-Apr-1996 #text_change 18-Jun-1999
C:Accession: B35094; A01544; A36052
Eur. J. Biochem. 15, 513-519, 1970
R:Kopin, A.S.; Wheeler, M.B.; Leiter, A.B.
Proc. Natl. Acad. Sci. U.S.A. 87, 2299-2303, 1990
A:Title: Secretin: structure of the precursor and tissue distribution of the mRNA.
A:Reference number: A35094; MUID:90192795
A:Accession: B35094
A:Molecule type: mRNA
A:Residues: 1-131 <KOP>
A:Cross-references: GB:M31496; NID:g164670; PIDN:AAA31121.1; PID:g164671
R:Mutt, V.; Jorpes, J.E.; Magnusson, S.
Eur. J. Biochem. 15, 513-519, 1970
A:Title: Structure of porcine secretin. The amino acid sequence.
A:Reference number: A91147; MUID:70282334
A:Accession: A01544
A:Molecule type: protein
A:Residues: 30-56 <MUT>
A:Note: tryptic peptides were sequenced
R:Gafvelin, G.; Joernvall, H.; Mutt, V.
Proc. Natl. Acad. Sci. U.S.A. 87, 6781-6785, 1990
A:Title: Processing of prosecretin: isolation of a secretin precursor from porcine intestines
A:Reference number: A36052; MUID:90370867
A:Accession: A36052
A>Status: preliminary
A:Molecule type: protein
A:Residues: 30-59, R, 92-131 <GAF>
R:Bocharsky, M.; Ondetti, M.A.; Levine, S.D.; Narayanan, V.L.; Saltza, M.V.; Sheehan, J.
Chem. Ind. 1966, 1757-1758, 1966
A:Title: Synthesis of a heptacosapeptide amide with the hormonal activity of secretin.
A:Reference number: A90916
A:Contents: annotation
A:Note: synthesis confirmed the proposed structure of the natural hormone
C:Superfamily: glucagon
C:Keywords: amidated carboxyl end; duodenal mucosa; duplication; hormone; secretagogue
F:1-18/Domain: signal sequence #status predicted <SIG>
F:30-56/Product: secretin #status experimental <MAT>
F:56/Modified site: amidated carboxyl end (Val) (amide in mature form from following gly

Query Match 2.3%; Score 7; DB 1; Length 131;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLQLQ 234

|||||||
Db 47 RLQRLQLQ 53

RESULT 30

JC2202
secretin precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 16-Jul-1999
C:Accession: JC2202; S34214
R:Lan, M.S.; Kajiyama, W.; Donadel, G.; Lu, J.; Notkins, A.L.
Biochem. Biophys. Res. Commun. 200, 1066-1071, 1994
A:Title: cDNA sequence and genomic organization of mouse secretin.
A:Reference number: JC2202; MUID:94234995
A:Accession: JC2202
A:Molecule type: mRNA
A:Residues: 1-133 <LAN>
A:Cross-references: EMBL:X73580; NID:g313710; PIDN:CAA51982.1; PID:g313711
C:Comment: This protein regulates the secretion of pancreatic juices and stimulates insulin
C:Superfamily: glucagon
C:Keywords: amidated carboxyl end; duplication; hormone; secretagogue
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-133/Product: prosecretin #status predicted <PRO>
F:32-58/Product: secretin #status predicted <MAT>
F:58/Modified site: amidated carboxyl end (Val) (amide in mature form from following

Query Match 2.3%; Score 7; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLQLQ 234

|||||||
Db 49 RLQRLQLQ 55

RESULT 31

A40959
secretin precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 20-Mar-1992 #sequence_revision 20-Mar-1992 #text_change 16-Jul-1999
C:Accession: A40886; A40959; A35094; A32544
R:Itoh, N.; Furuya, T.; Ozaki, K.; Ohta, M.; Kawasaki, T.
J. Biol. Chem. 266, 12595-12598, 1991
A:Title: The secretin precursor gene. Structure of the coding region and expression 1
A:Reference number: A40886; MUID:91286291
A:Accession: A40886
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-134 <ITO>
A:Cross-references: GB:M63984; NID:g206889; PIDN:AAA42127.1; PID:g206890
R:Kopin, A.S.; Wheeler, M.B.; Nishitani, J.; McBride, E.W.; Chang, T.; Chey, W.Y.; Le
Proc. Natl. Acad. Sci. U.S.A. 88, 5335-5339, 1991
A:Title: The secretin gene: evolutionary history, alternative splicing, and developme
A:Reference number: A40959; MUID:91271384
A:Accession: A40959
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-134 <KOP>
A:Cross-references: GB:M64033; NID:g206891; PIDN:AAA42128.1; PID:g206892
R:Kopin, A.S.; Wheeler, M.B.; Leiter, A.B.
Proc. Natl. Acad. Sci. U.S.A. 87, 2299-2303, 1990
A:Title: Secretin: structure of the precursor and tissue distribution of the mRNA.
A:Reference number: A35094; MUID:90192795
A:Accession: A35094
A>Status: preliminary
A:Molecule type: mRNA

R:Macri, C.
submitted to the EMBL Data Library, February 1994
A:Description: The sequence of *S. cerevisiae* cosmid 9177.
A:Reference number: S46671
A:Accession: S48994
A:Molecule type: DNA
A:Residues: 1-198 <MAC>
A:Cross-references: EMBL:U00029; MIPS:YHR213w; NID:g551322; PIDN:AAB69730.1; PID:g458918
R:Bussey, H.; Keng, T.; Storms, R.K.; Vo, D.; Zhong, W.; Fortin, N.; Barton, A.B.; Kabac
submitted to the EMBL Data Library, February 1994
A:Description: Sequencing of chromosome I of *Saccharomyces cerevisiae*: analysis of the
A:Reference number: S53458
A:Accession: S53471
A:Molecule type: DNA
A:Residues: 1-198 <BUS>
A:Cross-references: EMBL:L28920; MIPS:YAR062w; NID:g1616966; PIDN:AC09503.1; PID:g45615
A:Genetics: CHRI
C:Genetics: <CHR8>
A:Map position: 8R
C:Genetics: <CHR1>
A:Map position: 1R

Query Match 2.3%; Score 7; DB 2; Length 198;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 161 SSSSEQ 167
|||||||
Db 185 SSSSEQ 191

RESULT 42
G69906
conserved hypothetical protein yojJ - *Bacillus subtilis*
C:Species: *Bacillus subtilis*
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C:Accession: G69906
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Ehrlich, S.D.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel
Rieger, M.; Rivoita, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MUID:98044033
A:Accession: G69906
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-201 <KUN>
A:Cross-references: GB:Z99114; GB:AL009126; NID:g2634230; PIDN:CAB13835.1; PID:e1185415;
A:Experimental source: strain 168
C:Genetics:
A:Gene: yojJ

Query Match 2.3%; Score 7; DB 2; Length 201;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 269 DGALLVR 275
|||||||
Db 135 DGALLVR 141

RESULT 43
C75041
orotidine-5'-monophosphate decarboxylase PAB1505 - *Pyrococcus abyssi* (strain Orsay)
C:Species: *Pyrococcus abyssi*
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: C75041
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome s
A:Reference number: A75001
A:Accession: C75041
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-208 <KAW>
A:Cross-references: GB:AJ248287; GB:AL096836; NID:g5458657; PIDN:CAB50224.1; PID:g545
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB1505
C:Superfamily: hypothetical protein MJ0252

Query Match 2.3%; Score 7; DB 2; Length 208;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 137 GVIAPGT 143
|||||||
Db 135 GVIAPGT 141

RESULT 44
G83082
hypothetical protein PA4507 [imported] - *Pseudomonas aeruginosa* (strain PA01)
C:Species: *Pseudomonas aeruginosa*
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 15-Sep-2000
C:Accession: G83082
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
adman, S.; Yoon, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pa
A:Reference number: A82950
A:Accession: G83082
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-210 <STO>
A:Cross-references: GB:AE004864; GB:AE004091; NID:g9950740; PIDN:AAG07895.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA4507

Query Match 2.3%; Score 7; DB 2; Length 210;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 175 TALGLAL 181
|||||||
Db 39 TALGLAL 45

RESULT 45
S34163
homeotic protein Hoxd-3 - mouse
C:Species: *Mus musculus* (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999
C:Accession: S34163
R:Bedford, M.; Petr, L.
submitted to the EMBL Data Library, June 1993
A:Reference number: S34163
A:Accession: S34163
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-215 <BED>
A:Cross-references: EMBL:X73572; NID:g313248; PIDN:CAA51974.1; PID:g313249
C:Genetics:
A:Introns: 165/3

Query Match 2.3%; Score 7; DB 2; Length 215;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
|||||||

Db 133 SASSSSS 139

RESULT 46

JC7219
nuclear protein SR-25 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 21-Jul-2000
C:Accession: JC7219
R:Sasahara, K.; Yamaoka, T.; Moritani, M.; Tanaka, M.; Iwahana, H.; Yoshimoto, K.; Miyaguchi, K.; Biochem. Biophys. Res. Commun. 269, 444-450, 2000
A:Title: Molecular cloning and expression analysis of a putative nuclear protein, SR-25.
A:Reference number: JC7219; MUID:20175222
A:Accession: JC7219
A:Molecule type: mRNA
A:Residues: 1-229 <SAS>
A:Cross-references: DDBJ:AB035383; NID:g7619895; PIDN:BAA94743.1; PID:g7619896
A:Experimental source: MIN6 cell line
C:Comment: This protein is a highly hydrophilic nuclear protein with a serine-arginine rich A splicing factors.
C:Keywords: nucleus; RNA processing

Query Match 2.3%; Score 7; DB 2; Length 229;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
|||||||

Db 80 SASSSSS 86

RESULT 47

T24789
hypothetical protein T10C6.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T24789
R:McMurray, A.
submitted to the EMBL Data Library, March 1997
A:Reference number: Z19936
A:Accession: T24789
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-230 <WIL>
A:Cross-references: EMBL:Z93388; PIDN:CAB07655.1; GSPDB:GN00023; CESP:T10C6.5
A:Experimental source: clone T10C6
C:Genetics:
A:Gene: CESP:T10C6.5
A:Map position: 5
A:Introns: 57/3

Query Match 2.3%; Score 7; DB 2; Length 230;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
|||||||

Db 79 SASSSSS 85

RESULT 48

A48383
alpha s2-casein - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 03-May-1996
C:Accession: A48383
R:Alexander, L.J.; Das Gupta, N.A.; Beattie, C.W.
Anim. Genet. 23, 365-367, 1992
A:Title: The sequence of porcine alpha s2-casein cDNA.
A:Reference number: A48383; MUID:92367960
A:Accession: A48383
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-235 <ALE>
A:Experimental source: mammary gland
A:Note: sequence inconsistent with the nucleotide translation
A:Note: sequence extracted from NCBI backbone (NCBIN:110884, NCBIP:110885)
C:Superfamily: alpha-s2-casein

Query Match 2.3%; Score 7; DB 2; Length 235;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 ASSSSSE 166

|||||||

Db 69 ASSSSSE 75

RESULT 49

D71120
hypothetical protein PH0731 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 20-Jun-2000
C:Accession: D71120
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Se M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Ogo DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic A:Reference number: A71000; MUID:98344137
A:Accession: D71120
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-241 <KAW>
A:Cross-references: NID:g3236130; PIDN:BAA29822.1; PID:g3257139
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBa C:Genetics:
A:Gene: PH0731
C:Superfamily: hypothetical protein MJ0252

Query Match 2.3%; Score 7; DB 2; Length 241;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 GVIAPGT 143

|||||||

Db 168 GVIAPGT 174

RESULT 50

T05498
hypothetical protein T19K4.190 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
C:Accession: T05498
R:Bevan, M.; Wedler, H.; Wambutt, R.; Hoheisel, J.; Mewes, H.W.; Mayer, K.F.X.; Schue submitted to the Protein Sequence Database, April 1998
A:Reference number: Z15418
A:Accession: T05498

A;Molecule type: DNA
A;Residues: 1-272 <BEV>
A;Cross-references: EMBL:AL022373
A;Experimental source: cultivar Columbia; BAC clone T19K4
C;Genetics:
A;Map position: 4
A;Introns: 32/2; 72/1; 114/3; 150/3
A;Note: T19K4.190

Query Match 2.3%; Score 7; DB 2; Length 272;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
|||||||
DB 20 SASSSSS 26

Search completed: January 30, 2001, 17:07:54
Job time: 758 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 17:06:56 ; Search time 17.99 Seconds
(without alignments)
538.534 Million cell updates/sec

Title: US-09-518-931-2
Perfect score: 300
Sequence: 1 MRALEPGSLLCGLVLPALPA.....RVARMPLGERSVREPLPVH 300

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 88757 seqs, 32294092 residues

Word size : 6

Total number of hits satisfying chosen parameters: 938

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	8	2.7	130	1 OREX_MOUSE	O55241 mus musculus
2	8	2.7	130	1 OREX_RAT	O55232 rattus norv
3	8	2.7	131	1 OREX_HUMAN	O43612 homo sapien
4	8	2.7	131	1 OREX_PIG	O77668 sus scrofa
5	8	2.7	179	1 CAS2_RAT	P02667 rattus norv
6	8	2.7	184	1 CAS3_MOUSE	Q02862 mus musculus
7	8	2.7	494	1 ADRO_MOUSE	Q61578 mus musculus
8	8	2.7	530	1 YDEV_ECOLI	P77432 escherichia
9	8	2.7	626	1 GLMS_NOSS9	O68280 n glucosami
10	8	2.7	1172	1 TSP2_HUMAN	P35442 homo sapien
11	7	2.3	27	1 SECR_CANFA	P09910 canis famil
12	7	2.3	27	1 SECR_HUMAN	P09683 homo sapien
13	7	2.3	27	1 SECR_RABIT	P32647 oryctolagus
14	7	2.3	27	1 SECR_SHEEP	P31299 ovis aries
15	7	2.3	113	1 INM1_ECOLI	P33637 escherichia
16	7	2.3	113	1 INM1_ECOLI	P02985 escherichia
17	7	2.3	113	1 INM1_SHISO	P22558 shigella so
18	7	2.3	114	1 Y0J5_CAEEL	P34628 caenorhabdi
19	7	2.3	118	1 YNIF_AZOBR	P25316 azospirillu
20	7	2.3	131	1 SECR_PIG	P01279 sus scrofa
21	7	2.3	133	1 SECR_MOUSE	Q08535 mus musculus
22	7	2.3	134	1 SECR_RAT	P11384 rattus norv
23	7	2.3	145	1 TASM_BFDV	P13895 budgerigar
24	7	2.3	198	1 YH13_YEAST	P38896 saccharomyc
25	7	2.3	214	1 YDH1_HSV57	P25049 herpesvirus
26	7	2.3	230	1 RNS1_ARATH	P42813 arabidopsis
27	7	2.3	235	1 CAS2_PIG	P39036 sus scrofa
28	7	2.3	241	1 DCOP_PYRHO	O58462 pyrococcus
29	7	2.3	270	1 HIS7_ARATH	P34047 arabidopsis
30	7	2.3	323	1 FASA_BOVIN	P51867 bos taurus
31	7	2.3	323	1 Y495_MYCLE	P54579 mycobacteri
32	7	2.3	332	1 FASA_PIG	O77736 sus scrofa
33	7	2.3	334	1 SNC2_HUMAN	Q13487 homo sapien

34	7	2.3	337	1 NOD2_BRASN	P50328 bradyrhizob
35	7	2.3	353	1 HMZ1_DROME	P09089 drosophila
36	7	2.3	354	1 OTX1_HUMAN	P32242 homo sapien
37	7	2.3	355	1 OTX1_MOUSE	P80205 mus musculus
38	7	2.3	355	1 OTX1_RAT	Q63410 rattus norv
39	7	2.3	362	1 CKRA_HUMAN	P46092 homo sapien
40	7	2.3	401	1 YKO3_CAEEL	P34291 caenorhabdi
41	7	2.3	406	1 HEM1_AQUAE	O67314 aquifex aeo
42	7	2.3	406	1 SR40_YEAST	P32583 saccharomyc
43	7	2.3	410	1 HID_DROME	Q24106 drosophila
44	7	2.3	412	1 CLPX_AQUAE	O67356 aquifex aeo
45	7	2.3	417	1 HXD3_MOUSE	P09027 mus musculus
46	7	2.3	434	1 PEL_LILLO	P40973 lilium long
47	7	2.3	469	1 GSA_HORVU	P18492 hordeum vul
48	7	2.3	469	1 YS8_YEAST	P38248 saccharomyc
49	7	2.3	470	1 KLF4_HUMAN	O43474 homo sapien
50	7	2.3	474	1 KLF4_MOUSE	Q60793 mus musculus
51	7	2.3	479	1 OPRJ_PSEAE	Q51397 pseudomonas
52	7	2.3	482	1 RRB_MOUSE	P22605 mus musculus
53	7	2.3	487	1 HEP_DROME	Q23977 drosophila
54	7	2.3	487	1 TTDT_ECOLI	P39414 escherichia
55	7	2.3	494	1 AMY2_SALTY	P26613 salmonella
56	7	2.3	495	1 AMY2_ECOLI	P26612 escherichia
57	7	2.3	504	1 VLI_HPV51	P26536 human papil
58	7	2.3	535	1 MVIN_RHIME	P56882 rhizobium m
59	7	2.3	548	1 ERF_HUMAN	P50548 homo sapien
60	7	2.3	551	1 ERF_MOUSE	P70459 mus musculus
61	7	2.3	559	1 GAS1_YEAST	P22146 saccharomyc
62	7	2.3	587	1 TALA_BFDV	P13894 budgerigar
63	7	2.3	588	1 CAR7_CANAL	P43096 candida alb
64	7	2.3	600	1 SP96_DICDI	P14328 dictyosteli
65	7	2.3	606	1 ABD4_HUMAN	O14678 homo sapien
66	7	2.3	606	1 ABD4_MOUSE	O89016 mus musculus
67	7	2.3	636	1 YNR6_YEAST	P53882 saccharomyc
68	7	2.3	644	1 UL35_HCMVA	P16766 human cytom
69	7	2.3	744	1 BISC_RHOSH	Q24266 drosophila
70	7	2.3	754	1 RRG3_HUMAN	P54934 rhodobacter
71	7	2.3	801	1 PIPA_DICDI	Q02158 dictyosteli
72	7	2.3	810	1 GARN_DROME	P54362 drosophila
73	7	2.3	810	1 GARN_DROME	O64220 mycobacteri
74	7	2.3	836	1 VG26_BPMD2	P54447 drosophila
75	7	2.3	864	1 E78A_DROME	P97481 mus musculus
76	7	2.3	874	1 PAS1_MOUSE	O15265 homo sapien
77	7	2.3	892	1 ATX7_HUMAN	P54281 bos taurus
78	7	2.3	903	1 ECLC_BOVIN	P42762 arabidopsis
79	7	2.3	945	1 ERD1_ARATH	P48377 mus musculus
80	7	2.3	963	1 RFAL_MOUSE	P55060 homo sapien
81	7	2.3	971	1 CAS_HUMAN	O24600 arabidopsis
82	7	2.3	993	1 RPOT_ARATH	O24600 arabidopsis
83	7	2.3	1011	1 M3K6_HUMAN	O95382 homo sapien
84	7	2.3	1065	1 SED4_YEAST	P25365 saccharomyc
85	7	2.3	1075	1 FLO5_YEAST	P38894 saccharomyc
86	7	2.3	1093	1 AF17_HUMAN	P55198 homo sapien
87	7	2.3	1203	1 Y741_CAEEL	Q11069 caenorhabdi
88	7	2.3	1298	1 ICP4_HSV11	P08392 herpes simp
89	7	2.3	1322	1 YAG3_YEAST	P39712 saccharomyc
90	7	2.3	1324	1 MSH6_ARATH	O04716 arabidopsis
91	7	2.3	1417	1 BLM_HUMAN	P54132 homo sapien
92	7	2.3	1446	1 IE18_PRYKA	P34799 pseudorabie
93	7	2.3	1461	1 IE18_PRVIF	P11675 pseudorabie
94	7	2.3	2560	1 PPS2_BAGSU	P39846 bacillus su
95	7	2.3	3969	1 HRX_HUMAN	Q03164 homo sapien
96	7	2.3	4563	1 APB_HUMAN	P04114 homo sapien
97	6	2.0	47	1 V05_BPT7	P03777 bacterioph
98	6	2.0	61	1 XYLB_ACTMI	P12867 actinoplan
99	6	2.0	69	1 RUBL_BRAJA	P48344 bradyrhizob
100	6	2.0	70	1 COXQ_BOVIN	P10175 bos taurus
101	6	2.0	81	1 VSMF_CVFPF	P22655 porcine tra
102	6	2.0	82	1 VSMF_CVPRM	P24415 porcine res
103	6	2.0	82	1 YBDJ_ECOLI	P77506 escherichia
104	6	2.0	88	1 VG30_BPMD2	O64224 mycobacteri
105	6	2.0	90	1 Y060_NPVOP	O10317 orgyia pseu
106	6	2.0	92	1 HRK_MOUSE	P70678 mus musculus

107	1	92	1	LE10_HELAN	P46514	helianthus	180	6	2.0	199	1	YDHM_ECOLI	P76189	escherichia
108	1	98	1	FIXX_RHILT	P08710	rhizobium l	181	6	2.0	203	1	CTFI_RAT	P30086	rattus norv
109	6	102	1	YXCD_BACSU	P46334	bacillus su	182	6	2.0	203	1	IDI_MYCTU	P72002	mycobacteri
110	6	104	1	C552_BRAJA	P24059	bradyrhizob	183	6	2.0	203	1	YG3R_YEAST	P53288	saccharomyc
111	6	109	1	YQCC_ECOLI	P46919	escherichia	184	6	2.0	203	1	YG91_CAEEL	P34530	caenorhabdi
112	6	111	1	RL24_CHLTR	P28537	chlamydia t	185	6	2.0	205	1	GSCI_HUMAN	O15499	homo sapien
113	6	111	1	SC4_SCHCO	P16934	schizophtyll	186	6	2.0	206	1	CAS1_PIG	P39035	ovis aries
114	6	117	1	DHA_CRAPL	P22238	craterostig	187	6	2.0	206	1	CAS1_SHEEP	P04653	ovis aries
115	6	117	1	Y129_MYCPN	P75507	mycoplasma	188	6	2.0	207	1	CSF3_HUMAN	P09919	homo sapien
116	6	119	1	GVK2_HALSA	P33963	halobacteri	189	6	2.0	210	1	FTSQ_AGRU	O30990	agrobacteri
117	6	122	1	MPF2_PHLPR	P43214	phleum prat	190	6	2.0	213	1	HPI_DROVI	P29227	drosophila
118	6	122	1	SVS5_MOUSE	P30933	mus musculus	191	6	2.0	214	1	1432_SCHMA	Q26537	schistosoma
119	6	124	1	UROC_HUMAN	P55089	homo sapien	192	6	2.0	214	1	CAS1_CAPHI	P18626	capra hircu
120	6	128	1	DHX1_ARATH	P25863	arabidopsis	193	6	2.0	215	1	CALL_ARATH	P30188	arabidopsis
121	6	130	1	DH14_LYCES	P22240	lycopersico	194	6	2.0	216	1	YXCE_BACSU	P46335	bacillus su
122	6	137	1	MP15_LYMS	P25289	lymaea sta	195	6	2.0	218	1	COAT_CMVIX	Q66120	cucumber mo
123	6	138	1	MP15_LYMS	P31241	lymaea sta	196	6	2.0	218	1	COAT_CMVN	Q66138	cucumber mo
124	6	139	1	DH1_HORVO	P12951	hordeum vul	197	6	2.0	218	1	COAT_CMVY	P18027	cucumber mo
125	6	143	1	DH2_HORVU	P12952	hordeum vul	198	6	2.0	219	1	LIPB_SCHPO	O36017	schizosacch
126	6	144	1	RL16_BACHD	Q929K7	bacillus ha	199	6	2.0	220	1	1433_SPIOL	P29308	spinacia ol
127	6	145	1	DH11_GOSHI	P09442	gossypium h	200	6	2.0	220	1	ANTA_HYDMA	P38977	hydra magni
128	6	145	1	EXBB_HELPJ	Q92jp6	helicobacte	201	6	2.0	220	1	SER2_GALME	O96615	galleria me
129	6	145	1	EXBB_HELPY	O25897	helicobacte	202	6	2.0	223	1	GRPM_MOUSE	O88396	mus musculu
130	6	148	1	NIFZ_KLEPN	P08533	klebsiella	203	6	2.0	225	1	DH4_HORVU	P12949	hordeum vul
131	6	149	1	CYTM_HUMAN	Q15828	homo sapien	204	6	2.0	226	1	TSIS_SMSAV	P01128	simian sarc
132	6	149	1	DH15_WHEAT	Q00742	tritium ae	205	6	2.0	228	1	COX2_LUMTE	Q37545	lumbriacus t
133	6	150	1	CNRD_BOVIN	Q95142	bos taurus	206	6	2.0	228	1	DH25_ORYSA	P30287	oryza sativ
134	6	150	1	CNRD_CANFA	Q9xt54	canis famil	207	6	2.0	228	1	YDFH_ECOLI	P77577	escherichia
135	6	150	1	CNRD_HUMAN	O43924	homo sapien	208	6	2.0	231	1	GPHC_ALCEU	P40852	alcaligenes
136	6	150	1	CNRD_MOUSE	O55057	mus musculus	209	6	2.0	231	1	GPHP_ALCEU	P40853	alcaligenes
137	6	151	1	IN24_UROAP	P14777	uromyces ap	210	6	2.0	231	1	YC9G_SCHPO	Q09890	schizosacch
138	6	151	1	DH1D_ORYSA	P22913	oryza sativ	211	6	2.0	231	1	YJAH_ECOLI	P32681	escherichia
139	6	151	1	GLBT_CHITH	P02231	chironomus	212	6	2.0	235	1	1433_XENLA	P29309	xenopus lae
140	6	155	1	DHB_CRAPL	P22239	craterostig	213	6	2.0	236	1	1433_ENTHI	P42650	entamoeba h
141	6	155	1	ITRE_MAIZE	P01088	zea mays (m	214	6	2.0	236	1	GT6_SCHMA	P46435	schistosoma
142	6	155	1	Y120_YEREN	P27475	yersinia en	215	6	2.0	238	1	1432_ENTHI	P42649	entamoeba h
143	6	159	1	YK91_CAEEL	Q18268	caenorhabdi	216	6	2.0	238	1	RLJ_RICPR	Q92623	rickettsia
144	6	161	1	19KD_MYCAV	P46733	mycobacteri	217	6	2.0	238	1	YPE2_RHORI	P23139	rhodospirill
145	6	161	1	DH3_HORVU	P12948	hordeum vul	218	6	2.0	239	1	1431_ENTHI	P42648	entamoeba h
146	6	161	1	UL20_PRVN3	Q00702	pseudorabie	219	6	2.0	244	1	ARC3_CBCP	Q00901	clostridium
147	6	163	1	DH21_ORYSA	P12253	oryza sativ	220	6	2.0	244	1	FABG_VIBHA	P55336	vibrio harv
148	6	164	1	DH1E_ORYSA	P22911	oryza sativ	221	6	2.0	244	1	Y498_MYCLE	P54581	mycobacteri
149	6	164	1	DH1C_ORYSA	P22912	oryza sativ	222	6	2.0	245	1	143B_BOVIN	P29358	bos taurus
150	6	166	1	VPB_BPP2	P07696	bacterioph	223	6	2.0	245	1	143B_HUMAN	P31946	homo sapien
151	6	167	1	RL9_CHLMU	O46398	chlamydia m	224	6	2.0	245	1	143B_RAT	P35213	rattus norv
152	6	167	1	RL9_CHLTR	O46809	chlamydia t	225	6	2.0	245	1	143F_HUMAN	Q04917	homo sapien
153	6	168	1	DH1_MAIZE	P12950	zea mays (m	226	6	2.0	245	1	143F_MOUSE	P11576	mus musculu
154	6	168	1	VH11_AZOBR	P26723	azospirillu	227	6	2.0	245	1	143G_BOVIN	P29359	bos taurus
155	6	171	1	YC50_AQUAE	O67293	aquifex aeo	228	6	2.0	245	1	143T_HUMAN	P27348	homo sapien
156	6	172	1	YHL4_HCMVA	P09698	human cytom	229	6	2.0	245	1	143T_MOUSE	P35216	mus musculu
157	6	175	1	IF3_CHLTR	O84840	chlamydia t	230	6	2.0	245	1	143Z_MOUSE	P29312	homo sapien
158	6	177	1	GPIX_HUMAN	P14770	homo sapien	231	6	2.0	245	1	143Z_MOUSE	P35215	mus musculu
159	6	178	1	ATPD_STRBO	O50156	streptococc	232	6	2.0	245	1	143Z_SHEEP	P29361	ovis aries
160	6	180	1	LACB_CAPHI	P02756	capra hircu	233	6	2.0	246	1	143G_HUMAN	Q9un99	homo sapien
161	6	180	1	LACB_HORSE	P02758	equus cabal	234	6	2.0	246	1	143G_RAT	P35214	rattus norv
162	6	180	1	LACB_SHEEP	P02757	ovis aries	235	6	2.0	247	1	143B_SOYBN	Q96451	glycine max
163	6	180	1	YJQ2_YEAST	P46999	saccharomyc	236	6	2.0	247	1	GTS2_DROME	P41043	drosophila
164	6	182	1	YAF2_CAEEL	P52881	caenorhabdi	237	6	2.0	247	1	US09_HCMVA	P09729	human cytom
165	6	184	1	DHLE_RAPSA	P21298	raphanus sa	238	6	2.0	248	1	1433_CAEEL	P41932	caenorhabdi
166	6	185	1	DH14_ARATH	P42763	arabidopsis	239	6	2.0	248	1	1434_CAEEL	Q20655	caenorhabdi
167	6	186	1	DH18_ARATH	P30185	arabidopsis	240	6	2.0	248	1	143K_ARATH	P48348	arabidopsis
168	6	186	1	GL19_ORYSA	P29835	oryza sativ	241	6	2.0	248	1	143L_ARATH	P48349	arabidopsis
169	6	189	1	YHNN_CAEEL	Q94402	caenorhabdi	242	6	2.0	248	1	143S_HUMAN	P31947	homo sapien
170	6	191	1	CLD7_RAT	Q92111	rattus norv	243	6	2.0	249	1	143Z_DROME	P29310	drosophila
171	6	191	1	LIPB_ECOLI	P30976	escherichia	244	6	2.0	249	1	1431_LYCES	P93206	lycopersico
172	6	193	1	HIS7_SULSO	O33773	sulfolobus	245	6	2.0	249	1	143D_TOBAC	O49966	nicotiana t
173	6	197	1	HIS7_METJA	Q58109	methanococc	246	6	2.0	249	1	TPIS_TREPA	O83548	treponema p
174	6	197	1	Y146_NPVOP	O10375	orygia pseu	247	6	2.0	250	1	YL22_YEAST	Q07953	saccharomyc
175	6	198	1	C4_GIALA	Q01832	giardia lam	248	6	2.0	251	1	1433_FUCVE	Q39757	fucus vesic
176	6	198	1	CAS1_CAVPO	P04656	cavia porce	249	6	2.0	251	1	1433_TOBAC	Q41246	nicotiana t
177	6	198	1	GSCI_MOUSE	P56916	mus musculu	250	6	2.0	251	1	143R_ARATH	P42647	arabidopsis
178	6	198	1	YC69_PVRHO	O58984	pyrococcus	251	6	2.0	251	1	ARC3_CBDP	P15879	clostridium
179	6	199	1	RISA_CHLTR	O64410	chlamydia t	252	6	2.0	251	1	COBM_MYCTU	Q10672	mycobacteri

253	6	2.0	251	1	THIF_ECOLI	P30138	esch	326	6	2.0	285	1	ALDC_BACBR	P23616	bacillus br
254	6	2.0	252	1	1431_SCHMA	Q26540	schistosoma	327	6	2.0	286	1	3HAO_HUMAN	P46952	homo sapien
255	6	2.0	252	1	1433_DICDI	P34632	dichtosteli	328	6	2.0	286	1	3HAO_RAT	P46953	rattus norv
256	6	2.0	252	1	S910_SCHPO	Q06182	schizosacch	329	6	2.0	286	1	POTB_MYCPN	P75058	mycoplasma
257	6	2.0	254	1	1432_LYCPS	P93208	lycopersico	330	6	2.0	287	1	FCAR_HUMAN	P24071	homo sapien
258	6	2.0	254	1	1434_SOLFU	Q34643	solanum tub	331	6	2.0	287	1	RS2_MYCTU	Q10796	mycobacteri
259	6	2.0	254	1	143E_ARATH	P48347	arabidopsis	332	6	2.0	288	1	YG53_YEAST	P50091	saccharomyc
260	6	2.0	255	1	1435_LYCPS	P93210	lycopersico	333	6	2.0	289	1	NOLT_RHIFR	P33209	rhizobium f
261	6	2.0	255	1	143A_TOBAC	P93342	nicotiana t	334	6	2.0	289	1	NOLT_RHISN	P55714	rhizobium s
262	6	2.0	255	1	143B_TOBAC	Q49995	nicotiana t	335	6	2.0	291	1	TYE7_YEAST	P33122	saccharomyc
263	6	2.0	255	1	143E_HUMAN	P42655	homo sapien	336	6	2.0	292	1	YI25_MYCTU	Q50608	mycobacteri
264	6	2.0	255	1	143P_ARATH	P42644	arabidopsis	337	6	2.0	293	1	TYRC_ZYMMO	Q04983	zymomonas m
265	6	2.0	255	1	FECE_ECOLI	P15031	esch	338	6	2.0	294	1	END4_STRCO	Q9S2n2	streptomyce
266	6	2.0	257	1	143A_SOYBN	Q96450	glycine max	339	6	2.0	295	1	MYRA_MICGR	P37000	micromonosp
267	6	2.0	257	1	GLTL_HAEIN	P45022	haemophilus	340	6	2.0	296	1	Y495_MYCTU	Q11160	mycobacteri
268	6	2.0	257	1	TDXM_BOVIN	P35705	bos taurus	341	6	2.0	298	1	YC34_HELPY	O25832	helicobacte
269	6	2.0	257	1	YATD_CAEEL	P34462	caenorhabdi	342	6	2.0	298	1	YU90_YEAST	P47178	saccharomyc
270	6	2.0	258	1	1435_SOLFU	P93784	solanum tub	343	6	2.0	299	1	ICE1_SPOFR	P89116	spodoptera
271	6	2.0	258	1	1436_LYCPS	P93211	lycopersico	344	6	2.0	300	1	YP79_MYCTU	Q50642	mycobacteri
272	6	2.0	258	1	143C_SOYBN	Q96452	glycine max	345	6	2.0	301	1	FPG_RHIME	O59752	rhizobium m
273	6	2.0	258	1	143F_TOBAC	O49998	nicotiana t	346	6	2.0	301	1	PHLP_RAT	O63737	rattus norv
274	6	2.0	258	1	UL24_VZVD	P09288	varicella-z	347	6	2.0	303	1	CYSM_SALTY	P29848	salmonella
275	6	2.0	259	1	1433_CHLRE	P52908	chlamydomon	348	6	2.0	304	1	IGIR_PIG	Q29000	sus scrofa
276	6	2.0	259	1	1433_HELAN	O65352	helianthus	349	6	2.0	306	1	EC40_DAUCA	Q07322	daucus caro
277	6	2.0	259	1	1433_SOLFU	Q41418	solanum tub	350	6	2.0	306	1	OPPB_ECOLI	P31132	esch
278	6	2.0	259	1	1430_ARATH	Q01525	arabidopsis	351	6	2.0	306	1	OPPB_HAEIN	P45054	haemophilus
279	6	2.0	259	1	YG31_YEAST	P53269	saccharomyc	352	6	2.0	306	1	OPPB_SALTY	P48005	haemophilus
280	6	2.0	259	1	Y8FD_ECOLI	P45753	esch	353	6	2.0	308	1	CWEL_SCHPO	P53694	schizosacch
281	6	2.0	260	1	1433_LYCPS	P93209	lycopersico	354	6	2.0	308	1	PADI_STRCO	P41878	schizosacch
282	6	2.0	260	1	1433_OENHO	P29307	oenothera h	355	6	2.0	310	1	RS2_STRCO	O31212	streptomyce
283	6	2.0	260	1	1433_ORYSA	Q06967	oryza sativ	356	6	2.0	312	1	CAH4_HUMAN	P22748	homo sapien
284	6	2.0	260	1	1433_PEA	P46266	pisum sativ	357	6	2.0	312	1	DBHX_ANAPL	O57314	anas platyr
285	6	2.0	260	1	1434_LYCPS	P42652	lycopersico	358	6	2.0	314	1	VDGH_ECOLI	P76177	esch
286	6	2.0	260	1	143C_TOBAC	P93343	nicotiana t	359	6	2.0	315	1	SOL2_YEAST	P27262	saccharomyc
287	6	2.0	260	1	143E_DROME	P92177	drosophila	360	6	2.0	316	1	KLFP_RHOCA	P23386	rhodobacter
288	6	2.0	260	1	NRPN_RAT	O88780	rattus norv	361	6	2.0	317	1	RCN2_HUMAN	Q4257	homo sapien
289	6	2.0	261	1	1431_MAIZE	P49106	zea mays (m	362	6	2.0	318	1	NULM_BOVIN	P03887	bos taurus
290	6	2.0	261	1	1432_MAIZE	Q01526	zea mays (m	363	6	2.0	318	1	YIL7_YEAST	P40476	saccharomyc
291	6	2.0	261	1	143A_VICFA	P42653	vicia faba	364	6	2.0	319	1	PTHB_ECOLI	P56580	esch
292	6	2.0	261	1	143B_VICFA	P42654	vicia faba	365	6	2.0	319	1	RLUC_ECOLI	P23851	esch
293	6	2.0	261	1	143D_SOYBN	Q96453	glycine max	366	6	2.0	320	1	MDCF_RHIME	P50949	rhizobium m
294	6	2.0	261	1	COTN_BACSU	P54507	bacillus su	367	6	2.0	322	1	ANT_BPPI	P19655	bacterioph
295	6	2.0	261	1	VE39_NPVOP	Q05121	orgyia pseu	368	6	2.0	322	1	HYFC_ECOLI	P24193	esch
296	6	2.0	262	1	1433_TRIHA	Q99002	trichoderma	369	6	2.0	323	1	HFLC_BORBU	O51222	borrelia bu
297	6	2.0	262	1	143A_HORVU	P29305	hordeum vul	370	6	2.0	323	1	JUND_CHICK	P27921	gallus gall
298	6	2.0	262	1	143B_HORVU	Q43470	hordeum vul	371	6	2.0	325	1	YHU3_YEAST	P38844	saccharomyc
299	6	2.0	262	1	NODJ_RHIGA	P50333	rhizobium g	372	6	2.0	327	1	FASA_MOUSE	P25446	mus musculus
300	6	2.0	262	1	SPRE_RAT	P18297	rattus norv	373	6	2.0	329	1	BCSP_BRUAB	P12920	brucella ab
301	6	2.0	262	1	T2M1_MORSP	P11405	moraxella s	374	6	2.0	329	1	COAA_STRCO	O86779	streptomyce
302	6	2.0	263	1	143M_ARATH	Q96299	arabidopsis	375	6	2.0	330	1	GGPP_SULAC	P39464	sulfolobus
303	6	2.0	264	1	1433_CANAL	O42766	candida alb	376	6	2.0	331	1	MACS_BOVIN	P12624	bos taurus
304	6	2.0	264	1	1433_MESCR	P93259	mesembryant	377	6	2.0	331	1	NIXA_HELPY	Q48262	helicobacte
305	6	2.0	264	1	TRA2_DROME	P93018	drosophila	378	6	2.0	331	1	NOD2_BRAEL	P50324	bradyrhizob
306	6	2.0	265	1	143N_ARATH	Q96300	arabidopsis	379	6	2.0	332	1	OPT_HUMAN	Q9ubm4	homo sapien
307	6	2.0	265	1	VNST_RVFVZ	P21698	rift valley	380	6	2.0	333	1	PTHB_ERWAM	O32522	erwinia amy
308	6	2.0	266	1	1433_NEOCA	Q25538	neospora ca	381	6	2.0	333	1	SECF_RHOCA	O33518	rhodobacter
309	6	2.0	266	1	BMH1_YEAST	P29311	saccharomyc	382	6	2.0	334	1	LYTE_BACSU	P54421	bacillus su
310	6	2.0	267	1	143C_ARATH	P42643	arabidopsis	383	6	2.0	335	1	HXDD_HUMAN	P35453	homo sapien
311	6	2.0	267	1	143H_ARATH	P46077	arabidopsis	384	6	2.0	336	1	CAHC_ARATH	P27140	arabidopsis
312	6	2.0	267	1	CB23_PETSP	P40781	petunia sp.	385	6	2.0	336	1	PTHB_CLOBE	O32333	clostridium
313	6	2.0	268	1	1430_ARATH	P42645	arabidopsis	386	6	2.0	337	1	MREB_BACSU	Q01465	bacillus su
314	6	2.0	269	1	CB22_PEA	P07371	pisum sativ	387	6	2.0	337	1	YDTB_SCHPO	O14216	schizosacch
315	6	2.0	269	1	CHIB_MAIZE	P29023	zea mays (m	388	6	2.0	338	1	HMDX_METJA	Q58125	methanococ
316	6	2.0	270	1	RA24_SCHPO	P42656	schizosacch	389	6	2.0	338	1	RTCA_ECOLI	P46849	esch
317	6	2.0	270	1	RA25_SCHPO	P42657	schizosacch	390	6	2.0	339	1	HXDD_MOUSE	P70217	mus musculus
318	6	2.0	272	1	143E_TOBAC	O49997	nicotiana t	391	6	2.0	343	1	ILVE_HAEIN	P54689	haemophilus
319	6	2.0	272	1	BMH2_YEAST	P34730	saccharomyc	392	6	2.0	343	1	TRA_CORDI	P35879	corynebacte
320	6	2.0	273	1	GLV2_HAISA	P33964	halobacteri	393	6	2.0	344	1	ARC2_THETH	O50146	thermus aqu
321	6	2.0	275	1	VE39_NPVAC	P11042	autographa	394	6	2.0	347	1	UTR2_YEAST	P32623	saccharomyc
322	6	2.0	276	1	MURI_LACBR	P48797	lactobacilli	395	6	2.0	349	1	OPSD_MYRBE	P79798	myripristis
323	6	2.0	277	1	RS2_MYCLE	O33038	mycobacteri	396	6	2.0	349	1	OPSD_MYRBE	P79807	myripristis
324	6	2.0	283	1	TPM1_LOGMI	P31816	locusta mig	397	6	2.0	350	1	KAPA_MOUSE	P05132	mus musculus
325	6	2.0	284	1	CAS1_RAT	P02661	rattus norv	398	6	2.0	351	1	ROM1_BOVIN	P52205	bos taurus

399	6	2.0	351	1	ROM1_HUMAN	Q03395	homo sapien	472	6	2.0	409	1	EFTU_SYN7	P33171	synechococc
400	6	2.0	351	1	ROM1_MOUSE	P32958	mus musculus	473	6	2.0	410	1	HFLK_HAEN1	P44546	haemophilus
401	6	2.0	352	1	YK50_YEAST	P36151	saccharomyc	474	6	2.0	410	1	SEX3_MYCBO	O07129	mycobacteri
402	6	2.0	353	1	OPSD_SOLSO	Q99245	solea solea	475	6	2.0	410	1	SEX3_MYCTU	Q11155	mycobacteri
403	6	2.0	355	1	VALL_ABMVW	P21947	abutilon mo	476	6	2.0	410	1	YG31_MYCLE	O10178	mycobacteri
404	6	2.0	355	1	YDP6_SCHPO	O14009	schizosacch	477	6	2.0	414	1	NSR1_YEAST	P27476	saccharomyc
405	6	2.0	356	1	SPAS_SALTY	P40702	salmonella	478	6	2.0	416	1	EUT1_CAEEL	P28515	caenorhabdi
406	6	2.0	357	1	G6PT_HUMAN	P35575	homo sapien	479	6	2.0	416	1	PGK_AGABI	O94123	agaricus bi
407	6	2.0	357	1	G6PT_MOUSE	P35576	mus musculus	480	6	2.0	417	1	GA83_YEAST	Q04739	saccharomyc
408	6	2.0	357	1	G6PT_RAT	P34328	rattus norv	481	6	2.0	417	1	HS47_HUMAN	P29043	homo sapien
409	6	2.0	357	1	HRPN_BURSO	P35652	burkholderi	482	6	2.0	417	1	HS47_MOUSE	P19324	mus musculus
410	6	2.0	358	1	T2H2_HAEP4	P36433	haemophilus	483	6	2.0	417	1	HS47_RAT	P29457	rattus norv
411	6	2.0	360	1	UL29_HCMVA	P16764	human cytom	484	6	2.0	417	1	MAC1_YEAST	P35192	saccharomyc
412	6	2.0	360	1	Y871_RICPR	Q92C95	rickettsia	485	6	2.0	418	1	CBP2_HUMAN	P50454	homo sapien
413	6	2.0	361	1	PAX1_HUMAN	P15963	homo sapien	486	6	2.0	418	1	GSA_ARCFU	O29027	archaeoglob
414	6	2.0	361	1	VALL_PYMVU	P27258	potato yell	487	6	2.0	419	1	FAAA_HUMAN	P16930	homo sapien
415	6	2.0	361	1	VALL_TMOV	Q06657	tomato mott	488	6	2.0	420	1	IF2G_AERPE	O949c1	aeropyrum p
416	6	2.0	362	1	ERD1_YEAST	P16151	saccharomyc	489	6	2.0	424	1	CUS1_SOLTU	O04940	s phosphati
417	6	2.0	362	1	HUPK_BRAJA	P48342	bradyrhizob	490	6	2.0	424	1	NADB_MYCLE	Q49617	mycobacteri
418	6	2.0	362	1	PURK_HAEIN	P43850	haemophilus	491	6	2.0	424	1	KDTA_ECOLI	P23282	escherichia
419	6	2.0	362	1	ROAL_XENLA	P17130	xenopus lae	492	6	2.0	426	1	DHMH_PARVE	P23006	paracoccus
420	6	2.0	365	1	SUCC_METTH	O27115	methanobact	493	6	2.0	427	1	MYCN_SERCA	P26014	serinus can
421	6	2.0	367	1	BET3_MESAU	O09029	mesocricetu	494	6	2.0	429	1	PURK_MYCTU	P96881	mycobacteri
422	6	2.0	367	1	RMIL_AVILI	P10533	avian retro	495	6	2.0	430	1	TOLB_ECOLI	P19935	escherichia
423	6	2.0	368	1	HIP_RAT	P50503	rattus norv	496	6	2.0	431	1	P2X6_HUMAN	O15547	homo sapien
424	6	2.0	369	1	HIP_HUMAN	P50502	homo sapien	497	6	2.0	433	1	FTSZ_ARATH	Q42545	arabidopsis
425	6	2.0	370	1	GPRA_RAT	Q64121	rattus norv	498	6	2.0	438	1	UGPB_ECOLI	P10904	escherichia
426	6	2.0	370	1	MPK3_ARATH	Q39023	arabidopsis	499	6	2.0	439	1	COMP_BOVIN	P35445	bos taurus
427	6	2.0	375	1	ACT1_PNECA	P43239	pneumocysti	500	6	2.0	439	1	POTE_ECOLI	P24170	escherichia
428	6	2.0	375	1	ACT3_FUGRU	P53486	fugu rubrip	501	6	2.0	439	1	PURK_MYCLE	P46701	mycobacteri
429	6	2.0	376	1	ACTM_PISOC	P12171	pisaster oc	502	6	2.0	440	1	GAT4_MOUSE	O08369	mus musculus
430	6	2.0	376	1	YBC9_YEAST	P38201	saccharomyc	503	6	2.0	440	1	GAT4_RAT	P46152	rattus norv
431	6	2.0	377	1	TBX9_CAEEL	Q22289	caenorhabdi	504	6	2.0	441	1	GDF9_MOUSE	O07105	mus musculus
432	6	2.0	377	1	TDP_DROME	Q24318	drosophila	505	6	2.0	442	1	CYS4_DICDI	P54639	dictyosteli
433	6	2.0	381	1	ME2A_ASPFL	P46073	aspergillus	506	6	2.0	444	1	FXF2_HUMAN	Q12947	homo sapien
434	6	2.0	382	1	ACBC_ACTS5	Q92ae9	actinoplan	507	6	2.0	444	1	TIG_MYCPN	P75454	mycoplasma
435	6	2.0	384	1	DUS9_HUMAN	Q99556	homo sapien	508	6	2.0	445	1	AROA_CHLPN	Q926m0	chlamydia p
436	6	2.0	384	1	Y994_METTH	O27075	methanobact	509	6	2.0	445	1	YGCS_ECOLI	Q46909	escherichia
437	6	2.0	386	1	YG5V_YEAST	P53334	saccharomyc	510	6	2.0	446	1	BRNO_LACDL	P54104	lactobacilli
438	6	2.0	388	1	CHS_VIGUN	P51089	vigna ungu	511	6	2.0	447	1	YML3_ARATH	O23693	arabidopsis
439	6	2.0	388	1	CZF1_CANAL	P28875	candida alb	512	6	2.0	448	1	PRTE_ERWCH	P23597	erwinia chr
440	6	2.0	388	1	DACD_ECOLI	P33013	escherichia	513	6	2.0	450	1	RMIL_AVEVR	P27966	avian rous-
441	6	2.0	388	1	OXYR_MOUSE	P97926	mus musculus	514	6	2.0	450	1	YK22_YEAST	P36135	saccharomyc
442	6	2.0	388	1	OXYR_RAT	P70536	rattus norv	515	6	2.0	451	1	PPOX_MYCLE	Q50008	mycobacteri
443	6	2.0	391	1	HEM6_HORVU	Q42840	ordeum vul	516	6	2.0	451	1	YQB1_CAEEL	O09255	caenorhabdi
444	6	2.0	391	1	OXYR_SHEEP	Q28756	ovis aries	517	6	2.0	452	1	AMT_CORGL	P54146	corynebacte
445	6	2.0	392	1	GAG_BLVAV	P25058	bovine leuk	518	6	2.0	453	1	EUTB_ECOLI	P19635	escherichia
446	6	2.0	392	1	MFT1_YEAST	P33441	saccharomyc	519	6	2.0	453	1	VG_DROME	Q26366	drosophila
447	6	2.0	392	1	ODD_DROME	P23803	drosophila	520	6	2.0	454	1	TDGG_ECOLI	Q26330	escherichia
448	6	2.0	393	1	IL1S_CERAE	Q29612	cercopithe	521	6	2.0	454	1	Y095_MYCPN	P75538	mycoplasma
449	6	2.0	395	1	ADMR_RAT	P43142	mus musculus	522	6	2.0	455	1	YB12_YEAST	P38288	saccharomyc
450	6	2.0	395	1	ADMR_RAT	P31392	rattus norv	523	6	2.0	459	1	STB1_YEAST	P42845	saccharomyc
451	6	2.0	395	1	CARP_SYNRA	P81214	syncephalas	524	6	2.0	460	1	ACM1_HUMAN	P11229	homo sapien
452	6	2.0	395	1	HMPA_ERWCH	Q47266	erwinia chr	525	6	2.0	460	1	ACM1_MACMU	P56489	macaca mula
453	6	2.0	396	1	ARRS_RANCA	P51478	rana catesb	526	6	2.0	460	1	ACM1_MOUSE	P12657	mus musculus
454	6	2.0	396	1	ARRS_RANPI	P51479	rana pipien	527	6	2.0	460	1	ACM1_PIG	P04761	mus scrofa
455	6	2.0	396	1	HMPA_ECOLI	P24232	escherichia	528	6	2.0	460	1	ACM1_RAT	P08482	rattus norv
456	6	2.0	396	1	HMPA_SALTY	P26353	salmonella	529	6	2.0	460	1	SUOX_CHICK	P07850	gallus gall
457	6	2.0	398	1	GCH2_BACSU	P17620	bacillus su	530	6	2.0	461	1	ARLY_SYN3	P73257	synechocyst
458	6	2.0	398	1	IL1S_HUMAN	P27930	homo sapien	531	6	2.0	461	1	SSL1_YEAST	Q04673	saccharomyc
459	6	2.0	398	1	MEGL_PSEPU	P13254	pseudomonas	532	6	2.0	462	1	ALGC_PSEAE	P26276	pseudomonas
460	6	2.0	400	1	PAII_MUSVI	P50449	muscula vis	533	6	2.0	462	1	AMT2_MYCTU	Q11056	mycobacteri
461	6	2.0	400	1	PCAF_PSEPU	Q51956	pseudomonas	534	6	2.0	464	1	CAP_DICDI	P54654	dictyosteli
462	6	2.0	402	1	PAII_BOVIN	P13909	bos taurus	535	6	2.0	465	1	MOT3_HUMAN	O15427	homo sapien
463	6	2.0	403	1	TYRP_ECOLI	P18199	escherichia	536	6	2.0	467	1	EUTA_ECOLI	P76551	escherichia
464	6	2.0	403	1	YCOA_SYN7	P42460	synechococc	537	6	2.0	467	1	EUTA_SALTY	Q92fv2	salmonella
465	6	2.0	406	1	EFTU_DEISP	P33168	delionema s	538	6	2.0	468	1	SYE_THETH	P27000	thermus aqu
466	6	2.0	406	1	TPSP_HUMAN	P05154	homo sapien	539	6	2.0	470	1	ESCA_DROME	P25932	drosophila
467	6	2.0	407	1	LMO6_HUMAN	O43900	homo sapien	540	6	2.0	471	1	ARDY_CHLRE	P22675	chlamydomon
468	6	2.0	407	1	WCAI_ECOLI	P32057	escherichia	541	6	2.0	471	1	DAX1_PIG	P79386	mus scrofa
469	6	2.0	408	1	GSPE_ERWCA	P31705	erwinia car	542	6	2.0	471	1	GUX2_TRIRE	P07987	trichoderma
470	6	2.0	408	1	VPAP_VZVD	P09274	varicella-z	543	6	2.0	471	1	LEU2_BUCAI	P56934	buchnera ap
471	6	2.0	409	1	EFTU_ANANI	P18668	anacystis n	544	6	2.0	471	1	RBL2_HYDMR	Q59460	hydrogenovi

545	6	2.0	471	1	YB08_YEAST	P38257	saccharomyc
546	6	2.0	472	1	ARAE_ECOLI	P09830	escherichia
547	6	2.0	472	1	ARAE_KLEOX	P45598	klebsiella
548	6	2.0	473	1	RBL1_RHOCA	P32740	rhodobacter
549	6	2.0	473	1	RBL1_THIFE	P28895	thiobacillu
550	6	2.0	473	1	RBL2_THIFE	P07087	thiobacillu
551	6	2.0	473	1	TED_HUMAN	O75949	homo sapien
552	6	2.0	474	1	RBL_PSEHY	Q51856	pseudomonas
553	6	2.0	475	1	PPAT_MOUSE	P37231	homo sapien
554	6	2.0	475	1	PPAT_HUMAN	P37238	mus musculu
555	6	2.0	475	1	PPAT_RABIT	O19052	oryctolagus
556	6	2.0	475	1	SIM1_YEAST	P10472	saccharomyc
557	6	2.0	476	1	PCHA_PSEAE	Q51508	pseudomonas
558	6	2.0	477	1	PPAT_XENLA	P37234	xenopus lae
559	6	2.0	477	1	Y098_MYCGE	P47344	mycoplasma
560	6	2.0	479	1	EFT1_SOYBN	Q43467	glycine max
561	6	2.0	483	1	PREG_NEUCR	O06712	neurospora
562	6	2.0	484	1	YOD0_YEAST	Q08193	saccharomyc
563	6	2.0	489	1	MDM2_MOUSE	P23804	mus musculu
564	6	2.0	491	1	ADRO_HUMAN	P22570	homo sapien
565	6	2.0	494	1	ADRO_RAT	P56522	rattus norv
566	6	2.0	494	1	PHY5_AYESA	P06595	avena sativ
567	6	2.0	494	1	SYE_PSEAE	O9xc16	pseudomonas
568	6	2.0	496	1	CAT3_MAIZE	P18123	zea mays (m
569	6	2.0	497	1	DTPT_LACHE	O07380	lactobacill
570	6	2.0	497	1	SYC_CHLTR	O84787	chlamydia t
571	6	2.0	497	1	YQK1_CAEEL	Q09285	caenorhabdi
572	6	2.0	498	1	IL14_HUMAN	P40222	homo sapien
573	6	2.0	500	1	GAR2_SCHPO	P41891	schizosacch
574	6	2.0	501	1	YM05_YEAST	Q04472	saccharomyc
575	6	2.0	502	1	CP83_ARATH	P48421	arabidopsis
576	6	2.0	503	1	YKR5_YEAST	P34240	saccharomyc
577	6	2.0	503	1	YN23_YEAST	P53832	saccharomyc
578	6	2.0	504	1	P1XL_AZOCA	P26489	azorhizobi
579	6	2.0	508	1	EGRI_RAT	P08154	rattus norv
580	6	2.0	508	1	YLL2_YEAST	Q12303	saccharomyc
581	6	2.0	509	1	YMT9_MTXVL	P08073	myxoma viru
582	6	2.0	510	1	ATPA_RHORI	P05036	rhodospiril
583	6	2.0	511	1	GUNE_PSEFL	P18126	pseudomonas
584	6	2.0	514	1	TUPI_CANAL	P56093	candida alb
585	6	2.0	516	1	LEGB_GOSHI	P09800	gossypium h
586	6	2.0	520	1	RP54_RHIET	P49989	rhizobium e
587	6	2.0	521	1	TP25_CAEEL	Q10934	caenorhabdi
588	6	2.0	524	1	CPV1_ICTPU	Q92111	ictalurus p
589	6	2.0	524	1	YML6_YEAST	Q03655	saccharomyc
590	6	2.0	526	1	VG6G_STGMA	P12647	sigma virus
591	6	2.0	528	1	ATPA_STRLI	P50001	streptomyce
592	6	2.0	528	1	PR12_YEAST	P20457	saccharomyc
593	6	2.0	530	1	DRTS_SOYBN	P24120	glycine max
594	6	2.0	530	1	FUSA_BURCE	P24126	burkholderi
595	6	2.0	531	1	ACHC_ACHFU	P35903	achatina fu
596	6	2.0	532	1	ICAI_HUMAN	P05362	homo sapien
597	6	2.0	532	1	ICP0_HSVEB	P28990	equine herp
598	6	2.0	532	1	IPAT_SHIFL	P18014	shigella fl
599	6	2.0	532	1	SPG7_DICDI	P22698	dictyosteli
600	6	2.0	533	1	EGRI_MOUSE	P08046	mus musculu
601	6	2.0	533	1	VMCB_LAMBD	P03710	bacterioph
602	6	2.0	536	1	ATPA_SCHPO	P24487	schizosacch
603	6	2.0	536	1	THIP_ECOLI	P31549	escherichia
604	6	2.0	538	1	YEN1_SCHPO	O13695	schizosacch
605	6	2.0	538	1	THIP_HAEIN	P44985	haemophilus
606	6	2.0	540	1	YMA5_CAEEL	P34450	caenorhabdi
607	6	2.0	542	1	MASY_NEUCR	P28345	neurospora
608	6	2.0	545	1	ATPA_RAT	P15999	rattus norv
609	6	2.0	545	1	ATPA_XENLA	P08428	xenopus lae
610	6	2.0	545	1	FGR_FSVGR	P00544	feline sarc
611	6	2.0	547	1	CP78_MAIZE	P48420	zea mays (m
612	6	2.0	551	1	ATPA_NEUCR	P37211	neurospora
613	6	2.0	551	1	YG1F_YEAST	P53214	saccharomyc
614	6	2.0	552	1	ATPA_DROME	P35381	drosophila
615	6	2.0	553	1	ATPO_BOVIN	P19483	bos taurus
616	6	2.0	553	1	ATPA_HUMAN	P25705	homo sapien
617	6	2.0	553	1	ATPA_MOUSE	Q03265	mus musculu

6	2.0	554	1	HYES_HUMAN	P34913	homo sapien
6	2.0	556	1	PMGI_RICCO	P35493	ricinus com
6	2.0	557	1	IMVD_BACSU	P51785	bacillus su
6	2.0	558	1	FLAJ_METVO	O59649	methanococc
6	2.0	559	1	PMGI_MESCR	O42908	mesembryant
6	2.0	561	1	SSRP_RAT	O49331	rattus norv
6	2.0	563	1	METB_ARATH	P35217	arabidopsi
6	2.0	568	1	AF9_HUMAN	P42568	homo sapien
6	2.0	568	1	DISC_DROME	P37392	drosophila
6	2.0	568	1	YE1G_SCHPO	O13880	schizosacch
6	2.0	570	1	UL25_EBV	P03233	epstein-barr
6	2.0	570	1	UL49_HCMTA	P16786	human cytom
6	2.0	571	1	AMPL_LCYES	Q10712	lycopersicon
6	2.0	574	1	IPAA_SHIFL	P18009	shigella fl
6	2.0	577	1	P2B1_DROME	P48456	drosophila
6	2.0	579	1	M3K7_MOUSE	Q62073	mus musculu
6	2.0	581	1	IKR7_RAT	Q64716	rattus norv
6	2.0	583	1	FRDA_MYCTU	Q10760	mycobacteri
6	2.0	583	1	GTBI_MOUSE	O08582	mus musculu
6	2.0	584	1	GTBI_HUMAN	O00178	homo sapien
6	2.0	584	1	HMEN_DROVI	P09145	drosophila
6	2.0	587	1	LACP_KULJA	P07921	kluveromyc
6	2.0	588	1	VCLB_GOSHI	P09801	gossypium h
6	2.0	590	1	HMAA_DROME	P29555	drosophila
6	2.0	591	1	PU91_YEAST	P54113	s bifunctio
6	2.0	591	1	YAV3_XANCV	P14729	xanthomonas
6	2.0	592	1	CEA_CITFR	P04480	citrobacter
6	2.0	592	1	PTL1_ALCEU	P23536	s alcaligenes
6	2.0	592	1	P092_YEAST	P38009	s bifunctio
6	2.0	594	1	NIFA_RHLSN	Q35206	rhizobium s
6	2.0	594	1	RYK_MOUSE	Q01887	mus musculu
6	2.0	595	1	RUBB_PEA	P08927	pisum sativ
6	2.0	600	1	Y719_METJA	Q58129	methanococc
6	2.0	602	1	TX15_MOUSE	O73036	mus musculu
6	2.0	604	1	RYK_HUMAN	P34925	homo sapien
6	2.0	606	1	M3K7_HUMAN	O43318	homo sapien
6	2.0	608	1	GLMS_ECOLI	P17169	e glucosami
6	2.0	614	1	NBTB_RAT	P48056	rattus norv
6	2.0	617	1	VGF_RAT	P20156	rattus norv
6	2.0	619	1	VHPL_CAEEL	Q10038	caenorhabdi
6	2.0	622	1	NHSE_DROME	P51905	drosophila
6	2.0	624	1	YEXY_YEAST	P40090	saccharomyc
6	2.0	626	1	HRI_RABIT	P33279	oryctolagus
6	2.0	627	1	Y017_RICPR	Q93626	rickettsia
6	2.0	629	1	MYC0_MYCTU	O05809	mycobacteri
6	2.0	630	1	MUCL_MOUSE	Q02496	mus musculu
6	2.0	631	1	DHML_PARDE	P12293	paracoccus
6	2.0	631	1	ML16_HUMAN	Q14005	homo sapien
6	2.0	634	1	LE14_YEAST	P32389	saccharomyc
6	2.0	645	1	YFF4_CHICK	Q90980	gallus gall
6	2.0	646	1	YFF4_YEAST	P43549	saccharomyc
6	2.0	648	1	HEML_EMENI	P38092	emerella
6	2.0	651	1	CORO_YEAST	Q06440	saccharomyc
6	2.0	653	1	APPL_MOUSE	Q03157	mus musculu
6	2.0	655	1	YAGF_ECOLI	P77596	escherichia
6	2.0	658	1	PAK1_SCHPO	P50527	schizosacch
6	2.0	660	1	CCME_BRAJA	P45403	bradyrhizob
6	2.0	662	1	LOX2_BOVIN	P27479	bos taurus
6	2.0	662	1	LOX2_PIG	P16469	sus scrofa
6	2.0	662	1	LOXE_MOUSE	P35249	mus musculu
6	2.0	664	1	CNG2_RABIT	Q28718	oryctolagus
6	2.0	667	1	OS9_HUMAN	Q13438	homo sapien
6	2.0	669	1	SSRP_CHICK	Q04678	gallus gall
6	2.0	670	1	YER8_YEAST	P40059	saccharomyc
6	2.0	671	1	YPF1_ECOLI	P76562	escherichia
6	2.0	675	1	BMX_HUMAN	P51813	homo sapien
6	2.0	679	1	YG16_YEAST	P53070	saccharomyc
6	2.0	680	1	DCP_SALTY	P27236	salmonella
6	2.0	684	1	EP84_HCMTA	P17151	human cytom
6	2.0	685	1	YPCA_BACSU	P37483	bacillus su
6	2.0	686	1	PLB3_YEAST	Q08108	saccharomyc
6	2.0	690	1	HRPO_BURSO	P35656	burkholderi
6	2.0	690	1	IF2_THEMA	Q9wzn	thermotoga

691	6	2.0	695	1	MDL1_YEAST	P33310	saccharomyc	764	6	2.0	919	1	EXL3_HUMAN	O43909	homo sapien
692	6	2.0	704	1	NP14_RAT	P41777	rattus norv	765	6	2.0	928	1	KKR1_YEAST	P36003	saccharomyc
693	6	2.0	708	1	SRP_MOUSE	O08943	mus musculus	766	6	2.0	932	1	PMS1_HUMAN	P54277	homo sapien
694	6	2.0	708	1	VF40_HCMVA	P16753	human cytom	767	6	2.0	936	1	PHL1_YEAST	P39521	saccharomyc
695	6	2.0	709	1	SRP_HUMAN	Q08945	homo sapien	768	6	2.0	936	1	KSH4_HUMAN	O15457	homo sapien
696	6	2.0	711	1	MLLA_STRCO	Q53902	streptomyce	769	6	2.0	939	1	XPC_HUMAN	Q01831	homo sapien
697	6	2.0	712	1	IRAK_HUMAN	P51617	homo sapien	770	6	2.0	944	1	DNL4_YEAST	Q08387	saccharomyc
698	6	2.0	715	1	MX2_HUMAN	P20592	homo sapien	771	6	2.0	946	1	RHG4_HUMAN	P98171	homo sapien
699	6	2.0	720	1	ABI3_ARATH	O01593	arabidopsis	772	6	2.0	947	1	PMA2_YEAST	P19657	saccharomyc
700	6	2.0	720	1	NRG3_HUMAN	P56975	homo sapien	773	6	2.0	950	1	UPB1_USTMA	P40349	ustilago ma
701	6	2.0	725	1	TAP1_RAT	P36370	rattus norv	774	6	2.0	956	1	TSP3_HUMAN	P49746	homo sapien
702	6	2.0	730	1	COG9_MOUSE	P41245	mus musculus	775	6	2.0	956	1	TSP3_MOUSE	Q05895	mus musculus
703	6	2.0	731	1	DCOR_ECOLI	P21169	escherichia	776	6	2.0	958	1	AMYG_DEBOC	P22861	debaromyce
704	6	2.0	746	1	7UP2_DROME	P16376	drosophila	777	6	2.0	962	1	GUNA_PSEFL	P10476	pseudomonas
705	6	2.0	748	1	TAP1_HUMAN	Q03518	homo sapien	778	6	2.0	966	1	FIB1_PETMA	P02674	petromyzon
706	6	2.0	749	1	DNK3_SYNP7	P50022	synchococc	779	6	2.0	970	1	DA81_YEAST	P21657	saccharomyc
707	6	2.0	752	1	NEC1_RAT	P28840	rattus norv	780	6	2.0	976	1	FIBP_AD8B3	Q03553	bovine aden
708	6	2.0	753	1	NEC1_HUMAN	P29120	homo sapien	781	6	2.0	982	1	NIA_NEUR	P08619	neurospora
709	6	2.0	753	1	NEC1_MOUSE	P21662	mus musculus	782	6	2.0	989	1	SERA_PLAFG	P34754	mus musculus
710	6	2.0	755	1	COMP_RAT	P35444	rattus norv	783	6	2.0	993	1	EPB3_MOUSE	P54753	homo sapien
711	6	2.0	757	1	RTS1_YEAST	P38903	saccharomyc	784	6	2.0	998	1	EPB3_HUMAN	P29815	amsacta moo
712	6	2.0	759	1	YCD9_YEAST	P25569	saccharomyc	785	6	2.0	1002	1	SPHR_AMEPV	P34278	caenorhabdi
713	6	2.0	761	1	Y218_HUMAN	Q93075	homo sapien	786	6	2.0	1010	1	YKK1_CAEEL	O08644	mus musculus
714	6	2.0	764	1	CPT1_MOUSE	P97742	mus musculus	787	6	2.0	1014	1	EPB6_MOUSE	P41993	caenorhabdi
715	6	2.0	765	1	KRAB_HUMAN	P15056	homo sapien	788	6	2.0	1018	1	YKC2_CAEEL	Q04857	mus musculus
716	6	2.0	773	1	CPT1_HUMAN	P50416	homo sapien	789	6	2.0	1025	1	CA16_MOUSE	P42835	saccharomyc
717	6	2.0	773	1	CPT1_RAT	P32198	rattus norv	790	6	2.0	1041	1	EGT2_YEAST	Q64127	mus musculus
718	6	2.0	773	1	L119_DROME	Q24311	drosophila	791	6	2.0	1051	1	TF1A_MOUSE	P53836	saccharomyc
719	6	2.0	775	1	ICP0_HSV11	P08393	herpes simp	792	6	2.0	1060	1	YN18_YEAST	P39064	candida tsu
720	6	2.0	778	1	ACON_SCHPO	O13966	schizosacch	793	6	2.0	1070	1	AGLU_CANTIS	Q02308	drosophila
721	6	2.0	781	1	SP4_HUMAN	Q02446	homo sapien	794	6	2.0	1077	1	HLES_DROME	Q10251	schizosacch
722	6	2.0	788	1	BUD6_YEAST	P41697	saccharomyc	795	6	2.0	1079	1	IF2P_SCHPO	P38827	saccharomyc
723	6	2.0	789	1	ARNT_HUMAN	P27540	homo sapien	796	6	2.0	1080	1	SET1_YEAST	O96433	drosophila
724	6	2.0	789	1	YK25_CAEEL	P34332	caenorhabdi	797	6	2.0	1087	1	CCT_DROME	P35546	mus musculus
725	6	2.0	791	1	ARNT_MOUSE	P53762	mus musculus	798	6	2.0	1115	1	RET_MOUSE	O14248	schizosacch
726	6	2.0	794	1	SMO_CHICK	O42224	gallus gall	799	6	2.0	1125	1	YE62_SCHPO	P06593	avena sativ
727	6	2.0	798	1	ITB7_HUMAN	P26010	homo sapien	800	6	2.0	1128	1	PHY3_AVEA	P06594	avena sativ
728	6	2.0	802	1	PMT2_SCHPO	Q42832	schizosacch	801	6	2.0	1128	1	PHY4_AVEA	Q19272	caenorhabdi
729	6	2.0	805	1	E2F_DROME	Q27368	drosophila	802	6	2.0	1139	1	MSH5_CAEEL	P53692	schizosacch
730	6	2.0	806	1	ITB7_MOUSE	P26011	mus musculus	803	6	2.0	1140	1	RA18_SCHPO	O14432	homo sapien
731	6	2.0	806	1	RMIL_CHICK	O04982	gallus gall	804	6	2.0	1141	1	CN3A_HUMAN	Q62865	rattus norv
732	6	2.0	807	1	RMIL_COTJA	P34908	coturnix co	805	6	2.0	1152	1	CN3A_RAT	P06800	mus musculus
733	6	2.0	809	1	BCSB_ACEXY	P37717	acetobacter	806	6	2.0	1152	1	CD45_MOUSE	P05555	mus musculus
734	6	2.0	812	1	IF38_YEAST	P32497	saccharomyc	807	6	2.0	1153	1	ITAM_MOUSE	Q09531	caenorhabdi
735	6	2.0	813	1	YGR2_XENLA	Q03364	xenopus lae	808	6	2.0	1159	1	YQP4_CAEEL	Q92136	rattus norv
736	6	2.0	818	1	YBGQ_ECOLI	P75750	escherichia	809	6	2.0	1163	1	TSC1_RAT	Q92574	homo sapien
737	6	2.0	821	1	SPK1_YEAST	P22216	saccharomyc	810	6	2.0	1164	1	TSC1_HUMAN	P36170	saccharomyc
738	6	2.0	825	1	ICP0_HSV2H	P28284	herpes simp	811	6	2.0	1169	1	YK82_YEAST	Q95116	bos taurus
739	6	2.0	843	1	CYF1_BRUMA	Q27450	brugia mala	812	6	2.0	1170	1	TSP2_BOVIN	P35440	gallus gall
740	6	2.0	843	1	DPOL_HPBVJ	P17394	hepatitis b	813	6	2.0	1178	1	TSP1_CHICK	P54258	rattus norv
741	6	2.0	843	1	DPOL_HPBVO	P17395	hepatitis b	814	6	2.0	1183	1	DRPL_RAT	P34259	homo sapien
742	6	2.0	845	1	PPSA_AERPE	Q9yec5	aeropyrum p	815	6	2.0	1184	1	FBL2_HUMAN	P54259	homo sapien
743	6	2.0	845	1	TSO2_USTMA	Q99107	ustilago ma	816	6	2.0	1185	1	DRPL_HUMAN	P23226	drosophila
744	6	2.0	850	1	NRG2_HUMAN	O14511	homo sapien	817	6	2.0	1185	1	MAPX_DROME	P29597	homo sapien
745	6	2.0	854	1	UN33_CAEEL	Q01630	caenorhabdi	818	6	2.0	1187	1	TVK2_HUMAN	P29597	homo sapien
746	6	2.0	859	1	ALR1_YEAST	Q08269	saccharomyc	819	6	2.0	1189	1	YJH6_YEAST	P47035	saccharomyc
747	6	2.0	872	1	UL47_HSV4	P25073	equine herp	820	6	2.0	1196	1	DNBI_HSV2	P89452	herpes simp
748	6	2.0	877	1	PMT1_CANAL	O74189	candida alb	821	6	2.0	1197	1	DNBI_HSV2	P36384	herpes simp
749	6	2.0	881	1	YJH8_YEAST	P47033	saccharomyc	822	6	2.0	1199	1	N121_RAT	P52591	rattus norv
750	6	2.0	884	1	IMB_DROME	O18388	drosophila	823	6	2.0	1225	1	Y309_MYCGE	P47551	mycoplasma
751	6	2.0	886	1	ITH3_MOUSE	Q61704	mus musculus	824	6	2.0	1231	1	KIF4_MOUSE	P33174	mus musculus
752	6	2.0	887	1	ITH3_RAT	Q63416	rattus norv	825	6	2.0	1231	1	B3A2_RAT	P23347	rattus norv
753	6	2.0	887	1	MTS4_SCHPO	P87048	schizosacch	826	6	2.0	1236	1	POLS_WEEV	P13897	western equ
754	6	2.0	892	1	RA16_SCHPO	P36617	schizosacch	827	6	2.0	1237	1	B3A2_MOUSE	P13808	eastern equ
755	6	2.0	893	1	BOSS_DROVI	Q24738	drosophila	828	6	2.0	1239	1	POLS_EEVE	P08768	eastern equ
756	6	2.0	893	1	KNR_YEAST	P53739	saccharomyc	829	6	2.0	1240	1	B3A2_HUMAN	P04920	homo sapien
757	6	2.0	893	1	NIA_LEPMC	P36842	leptosphaer	830	6	2.0	1240	1	POLS_EEEV3	P27284	eastern equ
758	6	2.0	893	1	PMTX_SCHPO	O13898	schizosacch	831	6	2.0	1242	1	IRS1_HUMAN	P35568	homo sapien
759	6	2.0	898	1	KBF2_HUMAN	Q00653	homo sapien	832	6	2.0	1251	1	RBP2_PLAVB	Q00799	plasmodium
760	6	2.0	904	1	PMS1_YEAST	P14242	saccharomyc	833	6	2.0	1251	1	CD45_RAT	P04157	rattus norv
761	6	2.0	904	1	TORS_ECOLI	P39453	escherichia	834	6	2.0	1264	1	MOLR_ECOLI	P33345	escherichia
762	6	2.0	911	1	B3AT_HUMAN	P02730	homo sapien	835	6	2.0	1275	1	COBN_PSEDE	P29929	pseudomonas
763	6	2.0	915	1	SMOO_HUMAN	P53814	homo sapien	836	6	2.0	1288	1	OPLA_RAT	P97608	rattus norv

837	1	1290	6	2.0	1	P10844	clostridium
838	6	1297	6	2.0	1	P43847	haemophilus
839	6	1298	1	2.0	1	P35916	homo sapien
840	6	1300	1	2.0	1	Q01397	neurospora
841	6	1300	1	2.0	1	P14617	cavia porce
842	6	1301	1	2.0	1	P35832	drosophila
843	6	1304	1	2.0	1	P08575	homo sapien
844	6	1306	1	2.0	1	P32334	saccharomyc
845	6	1310	1	2.0	1	P09310	varicella-z
846	6	1317	1	2.0	1	Q10094	schizosacch
847	6	1363	1	2.0	1	P35917	mus musculus
848	6	1367	1	2.0	1	P08069	homo sapien
849	6	1390	1	2.0	1	Q93105	aedes aegypt
850	6	1403	1	2.0	1	O88322	mus musculus
851	6	1419	1	2.0	1	O13368	candida alb
852	6	1426	1	2.0	1	P16919	escherichia
853	6	1435	1	2.0	1	Q12766	homo sapien
854	6	1437	1	2.0	1	Q15440	homo sapien
855	6	1443	1	2.0	1	P30555	drosophila
856	6	1453	1	2.0	1	P30415	mus musculus
857	6	1462	1	2.0	1	P30414	homo sapien
858	6	1487	1	2.0	1	P28925	equine herp
859	6	1487	1	2.0	1	P17473	equine herp
860	6	1495	1	2.0	1	Q32333	homo sapien
861	6	1505	1	2.0	1	Q24167	drosophila
862	6	1522	1	2.0	1	Q05973	loliigo ble
863	6	1570	1	2.0	1	P54673	dictyosteli
864	6	1584	1	2.0	1	P47170	saccharomyc
865	6	1585	1	2.0	1	P34675	dictyosteli
866	6	1603	1	2.0	1	P35820	drosophila
867	6	1607	1	2.0	1	Q25410	lymnaea sta
868	6	1659	1	2.0	1	Q92093	oncorhynch
869	6	1677	1	2.0	1	Q90243	acipenser t
870	6	1677	1	2.0	1	Q98893	fundulus he
871	6	1690	1	2.0	1	P36252	thermotoga
872	6	1698	1	2.0	1	Q14999	homo sapien
873	6	1704	1	2.0	1	Q90508	fundulus he
874	6	1733	1	2.0	1	P04050	saccharomyc
875	6	1782	1	2.0	1	Q27309	bombyx mori
876	6	1807	1	2.0	1	P18709	xenopus lae
877	6	1823	1	2.0	1	Q91062	ichthyomyzo
878	6	1875	1	2.0	1	P16144	homo sapien
879	6	1882	1	2.0	1	P75109	mycoplasma
880	6	1900	1	2.0	1	P37297	saccharomyc
881	6	1905	1	2.0	1	P54683	dictyosteli
882	6	1912	1	2.0	1	P87498	gallus gall
883	6	1972	1	2.0	1	Q12888	homo sapien
884	6	2000	1	2.0	1	Q92794	homo sapien
885	6	2004	1	2.0	1	P04774	rattus norv
886	6	2009	1	2.0	1	P52362	herpes simp
887	6	2059	1	2.0	1	Q01513	podospora a
888	6	2145	1	2.0	1	Q01513	podospora a
889	6	2148	1	2.0	1	Q16927	aedes aegypt
890	6	2167	1	2.0	1	P39960	saccharomyc
891	6	2182	1	2.0	1	P08291	c genome po
892	6	2191	1	2.0	1	Q66474	e genome po
893	6	2201	1	2.0	1	Q95477	homo sapien
894	6	2201	1	2.0	1	P41233	mus musculus
895	6	2210	1	2.0	1	Q66802	ebola virus
896	6	2212	1	2.0	1	Q05318	ebola virus
897	6	2224	1	2.0	1	P12259	homo sapien
898	6	2301	1	2.0	1	P13899	t genome po
899	6	2303	1	2.0	1	P08544	t genome po
900	6	2327	1	2.0	1	O55017	mus musculus
901	6	2331	1	2.0	1	P5262	marburg vir
902	6	2336	1	2.0	1	Q02294	rattus norv
903	6	2339	1	2.0	1	Q00975	homo sapien
904	6	2339	1	2.0	1	Q05152	oryctolagus
905	6	2476	1	2.0	1	Q61687	mus musculus
906	6	2483	1	2.0	1	P3490	drosophila
907	6	2485	1	2.0	1	P36327	venezuelan
908	6	2490	1	2.0	1	Q12923	homo sapien
909	6	2512	1	2.0	1	P03317	sindbis vir
910	1	2554	6	2.0	1	7LES_DROME	
911	6	2647	1	2.0	1	ABP2_HUMAN	
912	6	2703	1	2.0	1	NOTC_DROME	
913	6	2805	1	2.0	1	MAPA_HUMAN	
914	6	2839	1	2.0	1	NFL_HUMAN	
915	6	2841	1	2.0	1	NFL_MOUSE	
916	6	2907	1	2.0	1	PN2_MOUSE	
917	6	2944	1	2.0	1	CA17_HUMAN	
918	6	2960	1	2.0	1	AHMK_HUMAN	
919	6	3005	1	2.0	1	ZFH2_DROME	
920	6	3033	1	2.0	1	POLG_HCVJB	
921	6	3144	1	2.0	1	VP13_YEAST	
922	6	3321	1	2.0	1	KEND_HUMAN	
923	6	3562	1	2.0	1	PGCV_CHICK	
924	6	3587	1	2.0	1	TYCB_BACER	
925	6	3649	1	2.0	1	ACVS_NOCLA	
926	6	3707	1	2.0	1	PGBM_MOUSE	
927	6	3726	1	2.0	1	TRX_DROME	
928	6	3801	1	2.0	1	LYST_HUMAN	
929	6	4289	1	2.0	1	TENX_HUMAN	
930	6	4344	1	2.0	1	DYHC_EMENT	
931	6	4349	1	2.0	1	DYHC_FUSSO	
932	6	4393	1	2.0	1	PGBM_HUMAN	
933	6	4543	1	2.0	1	LRP1_CHICK	
934	6	4544	1	2.0	1	LRP1_HUMAN	
935	6	4568	1	2.0	1	DYHC_CAPEL	
936	6	4687	1	2.0	1	PLEC_RAT	
937	6	4725	1	2.0	1	DYHC_DICDI	
938	6	5147	1	2.0	1	FAT_DROME	
939	6	5147	1	2.0	1	FAT_DROME	

ALIGNMENTS

RESULT 1

ID	OREX_MOUSE	STANDARD:	PRT:	130 AA.
AC	O55241:			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DE	OREXIN PRECURSOR (HYPCRETIN) (HCRT) [CONTAINS: OREXIN-A (HCRT1);			
DE	OREXIN-B (HCRT2)]			
GN	HCRT OR OX OR PPOX.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
RP	SEQUENCE FROM N.A.			
RN	MEDLINE=98150861; PubMed=9491897;			
RX	Sakurai T., Amemiya A., Ishii M., Matsuzaki I., Chemelli R.M.,			
RA	Tanaka H., Williams S.C., Richardson J.A., Kozlowski G.P., Wilson S.,			
RA	Arch J.R.S., Buckingham R.E., Haynes A.C., Carr S.A., Annan R.S.,			
RA	McNulty D.E., Liu W.-S., Terrett J.A., Elshourbagy N.A., Bergsma D.J.,			
RA	Yanagisawa M.;			
RT	"Orexins and orexin receptors: a family of hypothalamic neuropeptides			
RT	and G protein-coupled receptors that regulate feeding behavior.";			
RL	Cell 92:573-585(1998).			
RN	[2]			
SEQUENCE FROM N.A.				
STRAIN=C57BL/6J;				
RX	MEDLINE=98081872; PubMed=9419374;			
RA	de Lecea L., Kilduff T.S., Peyron C., Gao X.-B., Foye P.E.,			
RA	Danielson P.E., Fukuhara C., Battenberg E.L.F., Gautvik V.T.,			
RA	Bartlett P.S. II, Frankel W.N., van den Pol A.N., Bloom F.E.,			
RA	Gautvik K.M., Sutcliffe J.G.;			
RT	"The hypocretins: hypothalamus-specific peptides with neuroexcitatory			
RT	activity.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 95:322-327(1998).			
CC	-!- FUNCTION: REGULATORY PEPTIDES THAT MAY FUNCTION WITHIN THE CENTRAL			
CC	NERVOUS SYSTEM IN THE REGULATION OF FOOD INTAKE. BOTH OREXINS			
CC	STIMULATE FOOD CONSUMPTION IN A DOSE-DEPENDENT MANNER WHEN			
CC	ADMINISTERED INTRACEREBROVENTRICULARLY, THE EFFECTS OF OREXIN-B			

CC BEING LESSER-LASTING DUE EITHER TO A REDUCED RESISTANCE TO
 CC INACTIVATING PEPTIDASES, OR TO DIFFERENCES IN ITS BIOLOGICAL
 CC ACTION COMPARED TO THAT OF OREXIN-A. OREXIN-A BINDS BOTH TO OX1R
 CC AND OX2R WITH A HIGH AFFINITY, WHEREAS OREXIN-B BINDS ONLY TO OX2R
 CC WITH A SIMILAR HIGH AFFINITY.
 CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH PERIKARYAL ROUGH ENDOPLASMIC
 CC RETICULUM AS WELL AS CYTOPLASMIC LARGE GRANULAR VESICLES AT
 CC SYNAPSES (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: RESTRICTED TO NEURONAL CELL BODIES OF THE
 CC DORSAL AND LATERAL HYPOTHALAMUS.
 CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES AT PAIRED BASIC RESIDUES YIELD
 CC THE DIFFERENT ACTIVE PEPTIDES.
 CC -1- SIMILARITY: BELONGS TO THE OREXIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; AF041242; AAC40040.1; -;
 CC EMBL; AF019566; AAC02934.1; -;
 CC MGI:1202306; HCRT.
 CC INTERPRO: IPR001704; -;
 CC PRAM; PF02072; Orexin; 1.
 CC PRINTS; PR01091; OREXINPP.
 CC KW Neuropeptide; Cleavage on pair of basic residues; Signal; Amidation.
 CC FT SIGNAL 1 32 BY SIMILARITY.
 CC FT PEPTIDE 33 65 OREXIN-A.
 CC FT PEPTIDE 69 96 OREXIN-B.
 CC FT PROPEP 97 130
 CC FT MOD_RES 33 33 PYRROLIDONE CARBOXYLIC ACID (BY
 CC SIMILARITY).
 CC FT MOD_RES 65 65 AMIDATION (G-66 PROVIDE AMIDE GROUP) (BY
 CC SIMILARITY).
 CC FT MOD_RES 96 96 AMIDATION (G-97 PROVIDE AMIDE GROUP) (BY
 CC SIMILARITY).
 CC FT DISULFID 38 44 BY SIMILARITY.
 CC FT DISULFID 39 46 BY SIMILARITY.
 CC FT SEQUENCE 130 AA; 13503 MW; D3C223F9EE835F1C CRC64;
 CC SQ

Query Match 2.7%; Score 8; DB 1; Length 130;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 228 RLQRLQA 235
 |||||
 Db 78 RLQRLQA 85

RESULT 2
 OREX_RAT STANDARD; PRT; 130 AA.
 AC O5523;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE OREXIN PRECURSOR (HYPOCRETIN) (HCRT) [CONTAINS: OREXIN-A (HCRT1);
 DE OREXIN-B (HCRT2)].
 GN HCRT OR OX OR PPOX.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RN [1]
 RN SEQUENCE FROM N.A., AND SEQUENCE OF 33-65 AND 69-96.
 RP TISSUE-BRAIN;
 RC MEDLINE=98150861; PubMed=9491897;
 RX Sakurai T., Amemiya A., Ishii M., Matsuzaki I., Chemelli R.M.,
 RA Tanaka H., Williams S.C., Richardson J.A., Kozlowski G.P., Wilson S.,
 RA Arch J.R.S., Buckingham R.E., Haynes A.C., Carr S.A., Annan R.S.,

RA McNulty D.E., Liu W.-S., Terrett J.A., Elshourbagy N.A., Bergsma D.J.,
 RA Yanagisawa M.;
 RT "Orexins and orexin receptors: a family of hypothalamic neuropeptides
 RT and G protein-coupled receptors that regulate feeding behavior.";
 RL Cell 92:573-585(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RP STRAIN=SPRAGUE-DAWLEY;
 RX MEDLINE=98081872; PubMed=9419374;
 RA de Lecea L., Kilduff T.S., Peyron C., Gao X.-B., Foye P.E.,
 RA Danielson P.E., Fukuhara C., Battenberg E.L.F., Gautvik V.T.,
 RA Bartlett F.S. II, Frankel W.N., van den Pol A.N., Bloom F.E.,
 RA Gautvik K.M., Sutcliffe J.G.;
 RT "The hypocretins: hypothalamus-specific peptides with neuroexcitatory
 RT activity";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:322-327(1998).
 CC -1- FUNCTION: REGULATORY PEPTIDES THAT MAY FUNCTION WITHIN THE CENTRAL
 CC NERVOUS SYSTEM IN THE REGULATION OF FOOD INTAKE. BOTH OREXINS
 CC STIMULATE FOOD CONSUMPTION IN A DOSE-DEPENDENT MANNER WHEN
 CC ADMINISTERED INTRACEREBROVENTRICULARLY. THE EFFECTS OF OREXIN-B
 CC BEING LESSER-LASTING DUE EITHER TO A REDUCED RESISTANCE TO
 CC INACTIVATING PEPTIDASES, OR TO DIFFERENCES IN ITS BIOLOGICAL
 CC ACTION COMPARED TO THAT OF OREXIN-A. OREXIN-A BINDS BOTH TO OX1R
 CC AND OX2R WITH A HIGH AFFINITY, WHEREAS OREXIN-B BINDS ONLY TO OX2R
 CC WITH A SIMILAR HIGH AFFINITY.
 CC -1- SUBCELLULAR LOCATION: ASSOCIATED WITH PERIKARYAL ROUGH ENDOPLASMIC
 CC RETICULUM AS WELL AS CYTOPLASMIC LARGE GRANULAR VESICLES AT
 CC SYNAPSES.
 CC -1- TISSUE SPECIFICITY: EXCLUSIVELY DETECTED IN A DISCRETE SET IN
 CC HYPOTHALAMIC AND SUBTHALAMIC AREAS OF THE BRAIN AND, IN A SMALL
 CC AMOUNT, IN THE TESTIS. UNDETECTABLE IN NEURONS OF THE
 CC PARAVENTRICULAR, VENTROMEDIAL OR ARCuate NUCLEI, WHICH ARE KNOWN
 CC TO CONTAIN A VARIETY OF NEUROPEPTIDES ASSOCIATED WITH FOOD
 CC CONSUMPTION.
 CC -1- DEVELOPMENTAL STAGE: DETECTED AS EARLY AS EMBRYONIC DAY 18, BUT
 CC EXPRESSION INCREASED DRAMATICALLY AFTER THE THIRD POSTNATAL WEEK.
 CC -1- INDUCTION: BY NUTRITIONAL STATE, UP-REGULATED UNDER FASTED
 CC CONDITIONS.
 CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES AT PAIRED BASIC RESIDUES YIELD
 CC THE DIFFERENT ACTIVE PEPTIDES.
 CC -1- MASS SPECTROMETRY: MW=3558.7; MW_ERR=0.1; METHOD=MALDI;
 CC RANGE=33-65.
 CC -1- SIMILARITY: BELONGS TO THE OREXIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; AF041241; AAC40039.1; -;
 CC EMBL; AF019565; AAC02933.1; -;
 CC INTERPRO: IPR001704; -;
 CC PFAM; PF02072; Orexin; 1.
 CC PRINTS; PR01091; OREXINPP.
 CC KW Neuropeptide; Cleavage on pair of basic residues; Signal; Amidation.
 CC FT SIGNAL 1 32 OREXIN-A.
 CC FT PEPTIDE 33 65 OREXIN-B.
 CC FT PEPTIDE 69 96
 CC FT PROPEP 97 130
 CC FT MOD_RES 33 33 PYRROLIDONE CARBOXYLIC ACID.
 CC FT MOD_RES 65 65 AMIDATION (G-66 PROVIDE AMIDE GROUP).
 CC FT MOD_RES 96 96 AMIDATION (G-97 PROVIDE AMIDE GROUP).
 CC FT DISULFID 38 44
 CC FT DISULFID 39 46
 CC FT SEQUENCE 130 AA; 13645 MW; 00CAB259EDF2A404 CRC64;
 CC SQ

Query Match 2.7%; Score 8; DB 1; Length 130;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 228 RLRLQLQA 235
    |||||
    78 RLRLQLQA 85

RESULT 3
OREX_HUMAN
ID OREX_HUMAN STANDARD; PRT; 131 AA.
AC O43612;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE OREXIN PRECURSOR (HYPOCRETIN) (HCRT) [CONTAINS: OREXIN-A (HCRT1);
DE OREXIN-B (HCRT2)].
GN HCRT OR OX OR PPOX.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98150861; PubMed=9491897;
RA Sakurai T., Anemiyu A., Ishii M., Matsuzaki I., Chemelli R.M.,
RA Tanaka H., Williams S.C., Richardson J.A., Kozlowski G.P., Wilson S.,
RA Arch J.R.S., Buckingham R.E., Haynes A.C., Carr S.A., Annan R.S.,
RA McNulty D.E., Liu W.-S., Terrett J.A., Elshourbagy N.A., Bergsma D.J.,
RA Yanagisawa M.;
RT "Orexins and orexin receptors: a family of hypothalamic neuropeptides
RT and G protein-coupled receptors that regulate feeding behavior.";
RL Cell 92:573-585(1998).
CC -!- FUNCTION: REGULATORY PEPTIDES THAT MAY FUNCTION WITHIN THE CENTRAL
CC NERVOUS SYSTEM IN THE REGULATION OF FOOD INTAKE. BOTH OREXINS
CC STIMULATE FOOD CONSUMPTION IN A DOSE-DEPENDENT MANNER WHEN
CC ADMINISTERED INTRACEREBROVENTRICULARLY, THE EFFECTS OF OREXIN-B
CC BEING LESSER-LASTING DUE EITHER TO A REDUCED RESISTANCE TO
CC INACTIVATING PEPTIDASES, OR TO DIFFERENCES IN ITS BIOLOGICAL
CC ACTION COMPARED TO THAT OF OREXIN-A. OREXIN-A BINDS BOTH TO OX1R
CC AND OX2R WITH A HIGH AFFINITY, WHEREAS OREXIN-B BINDS ONLY TO OX2R
CC WITH A SIMILAR HIGH AFFINITY.
CC -!- SUBCELLULAR LOCATION: ASSOCIATED WITH PERIKARYAL ROUGH ENDOPLASMIC
CC RETICULUM AS WELL AS CYTOPLASMIC LARGE GRANULAR VESICLES AT
CC SYNAPSES (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN SUBTHALAMIC NUCLEUS
CC BUT UNDETECTABLE IN OTHER BRAIN REGIONS TESTED (HYPOTHALAMUS WAS
CC NOT TESTED) AND IN HEART, PLACENTA, LUNG, LIVER, SKELETAL MUSCLE,
CC KIDNEY AND PANCREAS.
CC -!- PTM: SPECIFIC ENZYMATIC CLEAVAGES AT PAIRED BASIC RESIDUES YIELD
CC THE DIFFERENT ACTIVE PEPTIDES.
CC -!- SIMILARITY: BELONGS TO THE OREXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF041240; AAC39600.1;
DR MIM; 602338;
DR INTERPRO; IPR001704;
DR PFAM; PF02072; Orexin; 1.
DR PRINTS; PR01091; OREXINPP.
KW Neuropeptide; Cleavage on pair of basic residues; Signal; Amidation.
FT SIGNAL 1 33 BY SIMILARITY.
FT PEPTIDE 34 66 OREXIN-A.
FT PEPTIDE 70 97 OREXIN-B.
FT PROPEP 98 131
FT MOD_RES 34 34 PYRROLIDONE CARBOXYLIC ACID (BY
FT SIMILARITY).
FT MOD_RES 66 66 AMIDATION (G-67 PROVIDE AMIDE GROUP) (BY
FT SIMILARITY).
FT MOD_RES 97 97 AMIDATION (G-98 PROVIDE AMIDE GROUP) (BY
SIMILARITY).

FT DISULFID 39 45 SIMILARITY).
FT DISULFID 40 47 BY SIMILARITY.
SQ SEQUENCE 131 AA; 13363 MW; 139D9C33E39E4EF1 CRC64;

Query Match 2.7%; Score 8; DB 1; Length 131;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLRLQLQA 235
    |||||
    79 RLRLQLQA 86

RESULT 4
OREX_PIG
ID OREX_PIG STANDARD; PRT; 131 AA.
AC O77668;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE OREXIN PRECURSOR (HYPOCRETIN) (HCRT) [CONTAINS: OREXIN-A (HCRT1);
DE OREXIN-B (HCRT2)].
GN HCRT OR OX OR PPOX.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN [1]
RP SEQUENCE FROM N.A., AND SYNTHESIS OF OREXIN-B.
RX TISSUE-HYPOTHALAMUS;
RX MEDLINE=99275384; PubMed=10343916;
RA Dyer C.J., Touchette K.J., Carroll J.A., Allee G.L., Matteri R.L.;
RT "Cloning of porcine prepro-orexin cDNA and effects of an intramuscular
RT injection of synthetic porcine orexin-B on feed intake in young
RT pigs.";
RL Domest. Anim. Endocrinol. 16:145-148(1999).
CC -!- FUNCTION: REGULATORY PEPTIDES THAT MAY FUNCTION WITHIN THE CENTRAL
CC NERVOUS SYSTEM IN THE REGULATION OF FOOD INTAKE. BOTH OREXINS
CC STIMULATE FOOD CONSUMPTION IN A DOSE-DEPENDENT MANNER WHEN
CC ADMINISTERED INTRACEREBROVENTRICULARLY, THE EFFECTS OF OREXIN-B
CC BEING LESSER-LASTING DUE EITHER TO A REDUCED RESISTANCE TO
CC INACTIVATING PEPTIDASES, OR TO DIFFERENCES IN ITS BIOLOGICAL
CC ACTION COMPARED TO THAT OF OREXIN-A. OREXIN-A BINDS BOTH TO OX1R
CC AND OX2R WITH A HIGH AFFINITY, WHEREAS OREXIN-B BINDS ONLY TO OX2R
CC WITH A SIMILAR HIGH AFFINITY (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: ASSOCIATED WITH PERIKARYAL ROUGH ENDOPLASMIC
CC RETICULUM AS WELL AS CYTOPLASMIC LARGE GRANULAR VESICLES AT
CC SYNAPSES (BY SIMILARITY).
CC -!- PTM: SPECIFIC ENZYMATIC CLEAVAGES AT PAIRED BASIC RESIDUES YIELD
CC THE DIFFERENT ACTIVE PEPTIDES.
CC -!- SIMILARITY: BELONGS TO THE OREXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF075241; AAC26827.1;
DR INTERPRO; IPR001704;
DR PFAM; PF02072; Orexin; 1.
DR PRINTS; PR01091; OREXINPP.
KW Neuropeptide; Cleavage on pair of basic residues; Signal; Amidation.
FT SIGNAL 1 33 BY SIMILARITY.
FT PEPTIDE 34 66 OREXIN-A.
FT PEPTIDE 70 97 OREXIN-B.
FT PROPEP 98 131
FT MOD_RES 34 34 PYRROLIDONE CARBOXYLIC ACID (BY
FT SIMILARITY).
FT MOD_RES 66 66 AMIDATION (G-67 PROVIDE AMIDE GROUP) (BY
FT SIMILARITY).

```



```

FT DISULFID 680 691 BY SIMILARITY.
FT CARBOHYD 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 316 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 457 457 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 584 584 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 710 710 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1069 1069 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1172 AA; 129955 MW; 2AC7BB230E4C6F5 CRC64;

Query Match 2.7%; Score 8; DB 1; Length 1172;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ALEGPGLS 10
DB 102 ALEGPGLS 109
|||||||

RESULT 11
SECR_CANFA STANDARD; PRT; 27 AA.
AC P09910;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE SECRETIN.
GN SCT.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Flissipedia; Canidae; Canis.
RN [1]
RP SEQUENCE.
RC TISSUE=INTESTINE;
RX MEDLINE=87314204; PubMed=3626755;
RA Shinomura Y., Eng J., Yalow R.S.;
RT "Dog secretin: sequence and biologic activity.";
RL Life Sci. 41:1243-1248(1987).
CC "1- FUNCTION: STIMULATES FORMATION OF NAHCO(3)-RICH PANCREATIC JUICE
AND SECRETION OF NAHCO(3)-RICH BILE AND INHIBITS HCL PRODUCTION
BY THE STOMACH.
CC "1- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
RN [1]
RP SEQUENCE.
RC TISSUE=SMALL INTESTINE;
RX MEDLINE=90259845; PubMed=2342988;
RA Gossen D., Buscall L., Cauvin A., Gourlet P., de Neef P., Rathe J.,
RA Robberecht P., Vandermeers-Piret M.C., Vandermeers A., Christophe J.;
RT "Amino acid sequence of VIP, PHI and secretin from the rabbit small
intestine.";
RL Peptides 11:123-128(1990).
CC "1- FUNCTION: STIMULATES FORMATION OF NAHCO(3)-RICH PANCREATIC JUICE
AND SECRETION OF NAHCO(3)-RICH BILE AND INHIBITS HCL PRODUCTION
BY THE STOMACH.
CC "1- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
DR PIR: C60415; C60415.
DR INTERPRO: IPR000532; -.
DR PFAM: PF00123; Hormone2; 1.
DR PRINTS: PR00275; GLUCAGON.
DR PROSITE: PS00260; GLUCAGON; 1.
KW Glucagon family; Hormone; Amidation.
FT MOD_RES 27 27
AMIDATION.
SQ SEQUENCE 27 AA; 3070 MW; 2D4015814F955B78 CRC64;

Query Match 2.3%; Score 7; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLQ 234
DB 18 RLQRLQ 24
|||||||

RESULT 12
SECR_HUMAN STANDARD; PRT; 27 AA.
AC P09683;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE SECRETIN.
GN SCT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE.
RA Carliquist M., Joernvall H., Forssmann W.-G., Thulin L., Johansson C.,
RA Mutt V.;
RT "Human secretin is not identical to the porcine/bovine hormone.";
RL IRCS Med. Sci. 13:217-218(1985).
CC "1- FUNCTION: STIMULATES FORMATION OF NAHCO(3)-RICH PANCREATIC JUICE
AND SECRETION OF NAHCO(3)-RICH BILE AND INHIBITS HCL PRODUCTION
BY THE STOMACH.
CC "1- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
DR PIR: S07443; S07443.
DR MIN: 182099; -.
DR INTERPRO: IPR000532; -.
DR PFAM: PF00123; hormone2; 1.
DR PRINTS: PR00275; GLUCAGON.
DR PROSITE: PS00260; GLUCAGON; 1.
KW Glucagon family; Hormone; Amidation.
FT MOD_RES 27 27
AMIDATION.
SQ SEQUENCE 27 AA; 3040 MW; 2D4015900B955B78 CRC64;

Query Match 2.3%; Score 7; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLQ 234
DB 18 RLQRLQ 24
|||||||

RESULT 13
SECR_RABIT STANDARD; PRT; 27 AA.
AC P32647;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE SECRETIN.
GN SCT.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
RN [1]
RP SEQUENCE.
RC TISSUE=SMALL INTESTINE;
RX MEDLINE=90259845; PubMed=2342988;
RA Gossen D., Buscall L., Cauvin A., Gourlet P., de Neef P., Rathe J.,
RA Robberecht P., Vandermeers-Piret M.C., Vandermeers A., Christophe J.;
RT "Amino acid sequence of VIP, PHI and secretin from the rabbit small
intestine.";
RL Peptides 11:123-128(1990).
CC "1- FUNCTION: STIMULATES FORMATION OF NAHCO(3)-RICH PANCREATIC JUICE
AND SECRETION OF NAHCO(3)-RICH BILE AND INHIBITS HCL PRODUCTION
BY THE STOMACH.
CC "1- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
DR PIR: C60415; C60415.
DR INTERPRO: IPR000532; -.
DR PFAM: PF00123; Hormone2; 1.
DR PRINTS: PR00275; GLUCAGON.
DR PROSITE: PS00260; GLUCAGON; 1.
KW Glucagon family; Hormone; Amidation.
FT MOD_RES 27 27
AMIDATION.
SQ SEQUENCE 27 AA; 3105 MW; 38A015800BDD3618 CRC64;

Query Match 2.3%; Score 7; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLQ 234
DB 18 RLQRLQ 24
|||||||

```

```
RESULT 14
SECR_SHEEP
ID SECR_SHEEP STANDARD; PRT; 27 AA.
AC P31299;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE SECRETIN.
GN SCT.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
RN [1]
RP SEQUENCE.
RC TISSUE=SMALL INTESTINE;
RX MEDLINE=91239834; PubMed=2034821;
RA Bounjoua Y., Vandermeers A., Robberecht P., Vandermeers-Piret M.C.,
RA Christophe J.;
RT "Purification and amino acid sequence of vasoactive intestinal
RT peptide, peptide histidine isoleucinamide and secretin from the ovine
RT small intestine."
RL Regul. Pept. 32:169-179(1991).
CC -!- FUNCTION: STIMULATES FORMATION OF NAHCO(3)-RICH PANCREATIC JUICE
CC AND SECRETION OF NAHCO(3)-RICH BILE AND INHIBITS HCL PRODUCTION
CC BY THE STOMACH.
CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
DR PIR; C60072; SESH.
DR HSP; P01274; IGCN.
DR INTERPRO; IPR00532; -.
DR PFAM; PF00123; hormone2; 1.
DR PRINTS; PR00275; GLUCAGON.
DR PROSITE; PS00260; GLUCAGON; 1.
DR GLUCAGON family; Hormone; Amidation.
KW MOD_RES 27 27 AMIDATION.
FT SEQUENCE 27 AA; 3056 MW; 2D4015814ED05B78 CRC64;

Query Match 2.3%; Score 7; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLQL 234
DB 18 RLQRLQL 24

RESULT 15
IMMO_ECOLI
ID IMMO_ECOLI STANDARD; PRT; 113 AA.
AC P33637;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE COLICIN E1 IMMUNITY PROTEIN (IMMEL) (MICROCIN E1 IMMUNITY PROTEIN).
GN IMM.
OS Escherichia coli.
OG Plasmid IncFII R100.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86319522; PubMed=3019092;
RA Ohtsubo H., Ryder T.B., Maeda Y., Armstrong K., Ohtsubo E.;
RT "DNA replication of the resistance plasmid R100 and its control."
RL Adv. Biophys. 21:115-133(1986).
CC -!- FUNCTION: THIS PROTEIN IS ABLE TO PROTECT A CELL, WHICH HARBORS
CC THE PLASMID COLEI ENCODING COLICIN E1, AGAINST COLICIN E1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M26840; AAA26071.1; -.
DR INTERPRO; IPR003061; -.
DR PRINTS; PR01298; MICROCIN.
KW Plasmid; Bacteriocin; Colicin.
SQ SEQUENCE 113 AA; 13369 MW; 484471D8C90B04C7 CRC64;

Query Match 2.3%; Score 7; DB 1; Length 113;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 LLCLVLA 17
DB 92 LLCLVLA 98

RESULT 16
IMM1_ECOLI
ID IMM1_ECOLI STANDARD; PRT; 113 AA.
AC P02985;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE COLICIN E1 IMMUNITY PROTEIN (IMMEL) (MICROCIN E1 IMMUNITY PROTEIN).
GN IMM.
OS Escherichia coli.
OG Plasmid COLEI.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86094231; PubMed=3936034;
RA Waleh N.S., Johnson P.H.;
RT "Structural and functional organization of the colicin E1 operon."
RL Proc. Natl. Acad. Sci. U.S.A. 82:8389-8393(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=80010893; PubMed=384144;
RA Oka A., Nomura N., Morita M., Sugisaki H., Sugimoto K.,
RA Takanami M.;
RT "Nucleotide sequence of small Colei derivatives: structure of the
RT regions essential for autonomous replication and colicin E1
RT immunity."
RL Mol. Gen. Genet. 172:151-159(1979).
CC -!- FUNCTION: THIS PROTEIN IS ABLE TO PROTECT A CELL, WHICH HARBORS
CC THE PLASMID COLEI ENCODING COLICIN E1, AGAINST COLICIN E1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J01566; AAB59122.1; -.
DR EMBL; M12543; AAA23066.1; -.
DR EMBL; V00268; CAA23529.1; -.
DR PIR; A03513; IMCELI.
DR INTERPRO; IPR003061; -.
DR PRINTS; PR01298; MICROCIN.
KW Plasmid; Bacteriocin; Colicin.
FT CONFLICT 105 105 E -> G (IN REF. 2).
SQ SEQUENCE 113 AA; 13359 MW; FFB8308B7C54E83E CRC64;

Query Match 2.3%; Score 7; DB 1; Length 113;
```

Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 LLCIVLA 17
Db 92 LLCIVLA 98
|||||

RESULT 17
IMM1_SHISO STANDARD; PRT: 113 AA.
AC P22558;
DT 01-AUG-1991 (Rel. 19, Created)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE COLICIN E1* IMMUNITY PROTEIN (IMMEL) (MICROCIN E1* IMMUNITY PROTEIN).
GN IMM.
OS Shigella sonnei.
OG Plasmid pKY-1.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Shigella.
RN [1]
RP SEQUENCE FROM N.A.
RA Higashi M., Hata M., Hase T., Yamaguchi K., Masamune Y.;
RT "The nucleotide sequence of cea and the region of origin of plasmid
pKY-1";
RL J. Gen. Appl. Microbiol. 32:433-442(1986).
CC - FUNCTION: THIS PROTEIN IS ABLE TO PROTECT A CELL, WHICH HARBORS
CC THE PLASMID pKY-1 ENCODING COLICIN E1*, AGAINST COLICIN E1*.
CC - SIMILARITY: 79% IDENTITY TO E-COLI COLICIN E1 IMMUNITY PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M37218; AAA98157.1; -
DR PIR; S11532; S11532.
DR INTERPRO; IPR003061; -
DR PRINTS; PR01298; MICROCIN.
KW Plasmid; Bacteriocin; Colicin.
SQ SEQUENCE 113 AA; 13423 MW; 89A4D21E69F34155 CRC64;

Query Match 2.3%; Score 7; DB 1; Length 113;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 LLCIVLA 17
Db 93 LLCIVLA 99
|||||

RESULT 18
YOU5_CAEEL STANDARD; PRT: 114 AA.
AC P34628;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE HYPOTHETICAL 13.0 KDA PROTEIN ZK353.5 IN CHROMOSOME III.
GN ZK353.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,

Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirksen J., Laistner N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkneen R.,
RA Sims M., Smaildon N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterson R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans";
RL Nature 368:32-38(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; LI5313; AAA28201.1; -
DR WORMPEP; ZK353.5; CE00389.
KW Hypothetical protein.
FT DOMAIN 11 18 SER-RICH
SQ SEQUENCE 114 AA; 13007 MW; A7F699F6F435989 CRC64;

Query Match 2.3%; Score 7; DB 1; Length 114;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
Db 11 SASSSSS 17
|||||

RESULT 19
YNIF_AZOBR STANDARD; PRT: 118 AA.
AC P25316;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE HYPOTHETICAL 12.3 KDA PROTEIN IN NIFK-NIFY INTERGENIC REGION.
OS Azospirillum brasilense.
OC Bacteria; Proteobacteria; alpha subdivision; Rhodospirillaceae;
OC Azospirillum.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92370074; PubMed=1823284;
RA Passaglia L.M.P., Nunes C.P., Zaha A., Schrank I.S.;
RT "The nifHDK operon in the free-living nitrogen-fixing bacteria
Azospirillum brasilense sequentially comprises genes H, D, K, an 353
bp orf and gene Y";
RL Braz. J. Med. Biol. Res. 24:649-675(1991).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M64344; AAB02345.1; -
DR PIR; S27476; S27476.
KW Hypothetical protein; Nitrogen fixation.
SQ SEQUENCE 118 AA; 12323 MW; 54783E0F3F4D8D53 CRC64;

Query Match 2.3%; Score 7; DB 1; Length 118;

Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSS 165
Db 73 SASSSS 79

RESULT 20
ID SECR_PIG STANDARD; PRT; 131 AA.
AC P01279;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE SECRETEIN PRECURSOR (FRAGMENT).
GN SCT.
OS Sus scrofa (Pig), Bos taurus (Bovine), and
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
[1]
RN SEQUENCE FROM N.A.
RP SPECIES-PIG;
RC MEDLINE-90192795; PubMed-2315322;
RA Kopin A.S., Wheeler M.B., Leiter A.B.;
RT "Secretin: structure of the precursor and tissue distribution of the
mRNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:2299-2303(1990).
RN [2]
RP SEQUENCE OF 30-56.
RC SPECIES-PIG;
RX MEDLINE-70282334; PubMed-5465996;
RA Mutt V., Jorpes J.E., Magnusson S.;
RT "Structure of porcine secretin. The amino acid sequence.";
RL Eur. J. Biochem. 15:513-519(1970).
RN [3]
RP SEQUENCE OF 30-59 AND 92-131.
RC SPECIES-PIG;
RX MEDLINE-90370867; PubMed-2395872;
RA Garvelin G., Joernvall H., Mutt V.;
RT "Processing of prosecretin: isolation of a secretin precursor from
porcine intestine.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:6781-6785(1990).
RN [4]
RP SYNTHESIS OF 30-131.
RC SPECIES-PIG;
RA Bodanszky M., Ondetti M.A., Levine S.D., Narayanan V.L., Saltza M.V.,
RA Sheehan J.T., Williams N.J., Sabo E.F.;
RT "Synthesis of a heptacosapeptide amide with the hormonal activity of
secretin.";
RL Chem. Ind. 1757-1758(1966).
RN [5]
RP SEQUENCE OF 30-56.
RC SPECIES-BOVINE;
RX MEDLINE-81237102; PubMed-7250377;
RA Carlquist M., Joernvall H., Mutt V.;
RT "Isolation and amino acid sequence of bovine secretin.";
RL FEBS Lett. 127:71-74(1981).
RN [6]
RP SEQUENCE OF 30-56.
RC SPECIES-C.PORCELLUS;
RX MEDLINE-90254163; PubMed-2340294;
RA Buscall L., Cauvin A., Gourlet P., Gossen D., de Neef P., Rathe J.,
RA Robberecht P., Vandermeers-Piret M.-C., Vandermeers A., Christophe J.;
RT "Purification and amino acid sequence of vasoactive intestinal
peptide, peptide histidine isoleucineamide (1-27) and secretin from
the small intestine of guinea pig.";
RL Biochim. Biophys. Acta 1038:355-359(1990).
RN [7]
RP STRUCTURE BY NMR OF SECRETIN.
RX MEDLINE-88151942; PubMed-2831051;
RA Clore G.M., Nilges M., Bruenger A., Gronenborn A.M.;

"Determination of the backbone conformation of secretin by restrained
molecular dynamics on the basis of interproton distance data.";
Eur. J. Biochem. 171:479-484(1988).
RN [8]
RP STRUCTURE BY NMR OF SECRETIN.
RX MEDLINE-87191017; PubMed-2883029;
RA Gronenborn A.M., Bovermann G., Clore G.M.;
RT "A 1H-NMR study of the solution conformation of secretin. Resonance
assignment and secondary structure.";
RL FEBS Lett. 215:88-94(1987).
CC -!- FUNCTION: STIMULATES FORMATION OF NAHCO(3)-RICH PANCREATIC JUICE
AND SECRETION OF NAHCO(3)-RICH BILE AND INHIBITS HCL PRODUCTION
BY THE STOMACH.
CC -!- PHARMACEUTICAL: AVAILABLE UNDER THE NAME SECRETIN-FERRING (FERRING
PHARMACEUTICALS).
CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; M31496; AAA311121.1;
DR PIR; A01544; SEPG.
DR PIR; A91291; SEBO.
DR PIR; A36052; A36052.
DR PIR; B35094; B35094.
DR HSSP; P01274; IGCN.
DR INTERPRO; IPR000532;
DR PFAM; PF00123; hormone2; 1.
DR PROSITE; PS00260; GLUCAGON; 1.
KW Glucagon family; Hormone; Amidation; Signal; Pharmaceutical.
FT NON_TER 1
FT SIGNAL <1 18
FT PEPTIDE 30 56
FT MOD_RES 56 56
SQ SEQUENCE 131 AA; 14277 MW; 1A24BDDA600E4E34 CRC64;

Query Match 2.3%; Score 7; DB 1; Length 131;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLQ 234
Db 47 RLQRLQ 53

RESULT 21
SECR_MOUSE
ID SECR_MOUSE STANDARD; PRT; 133 AA.
AC Q08535;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE SECRETEIN PRECURSOR.
GN SCT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RX SEQUENCE FROM N.A.
RX MEDLINE-94234995; PubMed-8179583;
RA Lan M.S., Kajiyama W., Donadel G., Lu J., Notkins A.L.;
RT "CDNA sequence and genomic organization of mouse secretin.";
RL Biochem. Biophys. Res. Commun. 200:1086-1071(1994).
CC -!- FUNCTION: STIMULATES FORMATION OF NAHCO(3)-RICH PANCREATIC JUICE
AND SECRETION OF NAHCO(3)-RICH BILE AND INHIBITS HCL PRODUCTION
BY THE STOMACH.
CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U07568; AAA18453.1; -
DR EMBL; X73580; CAA51982.1; -
DR PIR; JC2202; JC2202.
DR PIR; S34214; S34214.
DR MGI; 99466; SCT.
DR INTERPRO; IPR000532; -
DR PFAM; PF00123; hormone2; 1.
DR PRINTS; PR00275; GLUCAGON.
DR PROSITE; PS00260; GLUCAGON; 1.
KW Glucagon family; Hormone; Amidation; Signal.
FT SIGNAL 1 22 BY SIMILARITY.
FT PEPTIDE 32 58 SECRETIN (BY SIMILARITY).
FT MOD_RES 58 58 AMIDATION (G-59 PROVIDE AMIDE GROUP).
SQ SEQUENCE 133 AA; 14914 MW; 9869CBFC74CA9709 CRC64;

Query Match 2.3%; Score 7; DB 1; Length 133;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 LRLRLQ 234
   |||||
DB 49 LRLRLQ 55

RESULT 22
ID SECR_RAT STANDARD; PRT; 134 AA.
AC P11384;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE SECRETIN PRECURSOR.
GN SCT.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-90192795; PubMed-2315322;
RA Kopin A.S., Wheeler M.B., Leiter A.B.;
RT "Secretin: structure of the precursor and tissue distribution of the
RT mRNA."
RL Proc. Natl. Acad. Sci. U.S.A. 87:2299-2303(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-91271384; PubMed-1711228;
RA Kopin A.S., Wheeler M.B., Nishitani J., McBride E.W., Chang T.M.,
RA Chey W.Y., Leiter A.B.;
RT "The secretin gene: evolutionary history, alternative splicing, and
RT developmental regulation."
RL Proc. Natl. Acad. Sci. U.S.A. 88:5335-5339(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RX MEDLINE-91286291; PubMed-2061329;
RA Itoh N., Furuya T., Ozaki K., Kawasaki T.;
RT "The secretin precursor gene. Structure of the coding region and
RT expression in the brain."
RL J. Biol. Chem. 266:12595-12598(1991).
RN [4]
RP SEQUENCE OF 33-59.
RX MEDLINE-89246545; PubMed-2719704;
RA Gossen D., Vandermeers A., Vandermeers-Piret M.-C., Rathe J.,

```

```

RA Cauvin A., Robberecht P., Christophe J.;
RT "Isolation and primary structure of rat secretin.";
RL Biochem. Biophys. Res. Commun. 160:862-867(1989).
CC -!- FUNCTION: STIMULATES FORMATION OF NAHCO(3)-RICH PANCREATIC JUICE
CC AND SECRETION OF NAHCO(3)-RICH BILE AND INHIBITS HCL PRODUCTION
CC BY THE STOMACH.
CC -!- SIMILARITY: BELONGS TO THE GLUCAGON FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M31495; AAA42126.1; -
DR EMBL; M64033; AAA42128.1; -
DR EMBL; M63984; AAA42127.1; -
DR PIR; A35094; A35094.
DR PIR; A32544; A32544.
DR PIR; A40886; A40886.
DR PIR; A40959; A40959.
DR HSSP; P01274; 1GCM.
DR INTERPRO; IPR000532; -
DR PFAM; PF00123; hormone2; 1.
DR PRINTS; PR00275; GLUCAGON.
DR PROSITE; PS00260; GLUCAGON; 1.
KW Glucagon family; Hormone; Amidation; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT PEPTIDE 33 59 SECRETIN.
FT MOD_RES 59 59 AMIDATION (G-60 PROVIDE AMIDE GROUP).
SQ SEQUENCE 134 AA; 15072 MW; D9FA1A4C1F7C86E6 CRC64;

Query Match 2.3%; Score 7; DB 1; Length 134;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 LRLRLQ 234
   |||||
DB 50 LRLRLQ 56

RESULT 23
ID TASM_BFDV STANDARD; PRT; 145 AA.
AC P13895;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE SMALL T ANTIGEN.
OS Budgerigar fledgling disease virus (BFDV).
OC Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Polyomavirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-88265888; PubMed-2838972;
RA Rott O., Kroeger M., Mueller H., Hobom G.;
RT "The genome of budgerigar fledgling disease virus, an avian
RT polyomavirus."
RL Virology 165:74-86(1988).
RN [2]
RP SEQUENCE FROM N.A.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M20775; AAB59761.1; -
DR PIR; E29194; TVVPBD.
DR HSSP; P08622; 1XBL.

```

```

INTERPRO: IPRO01623; -.
PFAM: PF00226; DnaJ; 1.
PROSITE: PS50076; DNAJ_2; 1.
Early protein.
SEQUENCE 145 AA: 16988 MW; 794AC88B2223046D CRC64;

Query Match          2.3%; Score 7; DB 1; Length 145;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

259 RRTELL 265
|||||||
5 RRTELL 11

RESULT 24
YH13 YEAST
IID_YH13 YEAST STANDARD; PRT; 198 AA.
AC P38896;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL 21.1 KDA PROTEIN IN FLO5-PHO12 AND FLO1-PHO11 INTERGENIC
DE REGIONS.
DE YHR213W AND YAR062W.
GN Saccharomyces cerevisiae (Baker's yeast).
OS Saccharomycetes: Saccharomycetales;
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC Saccharomycetaceae; Saccharomycetes.
[1]
RN Saccharomycetes FROM N.A. (YHR213W).
RN RP
RN RC STRAIN=S288C / AB972;
RN RX MEDLINE=94378003; PubMed=8091229;
RN RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RN RA Du Z., Favello A., Fulton L., Gattung S., Geisel C., Kirsten J.,
RN RA Kucaba T., Hillier L., Jlier M., Johnston L., Langston Y.,
RN RA Latreille P., Louis E.J., Macri C., Mardis E., Meneses S., Mouser L.,
RN RA Nhan M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,
RN RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
RN RA Vaudin M.;
RN RA "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RN RT VIII.";
RN RL Science 265:2077-2082(1994).
[2]
RN Saccharomycetes FROM N.A. (YAR062W).
RN RP
RN RC STRAIN=S288C / AB972;
RN RX Bussey H., Keng T., Storms R.K., Vo D., Zhong W., Fortin N.,
RN RA Barton A.B., Kaback D.B., Clark M.W.;
RN RA Barton A.B., Kaback D.B., Clark M.W.;
RN RA Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
[3]
RN Saccharomycetes FROM N.A. (YAR062W).
RN RP
RN RC STRAIN=S288C / AB972;
RN RX MEDLINE=95249563; PubMed=7731989;
RN RA Bussey H., Kaback D.B., Zhong W., Vo D.T., Clark M.W., Fortin N.,
RN RA Hail J., Ouellette B.F., Keng T., Barton A.B., Su Y., Davies C.K.,
RN RA Storms R.K.;
RN RA "The nucleotide sequence of chromosome I from Saccharomyces
RN RT cerevisiae.";
RN RL Proc. Natl. Acad. Sci. U.S.A. 92:3809-3813(1995).
-1- SIMILARITY: TO YEAST PROTEIN FLO1.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
-----
CC EMBL; U00029; AAB69730.1; -.
CC EMBL; L28920; AAC09503.1; -.
CC PIR; S48994; S48994.
CC INTERPRO: IPR001389; -.

```


RA Bariola P.A., Howard C.J., Taylor C.B., Verburg M.T., Jaglan V.D.,
RT Green P.J.;
RL "The Arabidopsis ribonuclease gene RNS1 is tightly controlled in
response to phosphate limitation.";
RL Plant J. 6:673-685(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Rounsley S.D., Lin X., Kaul S., Shea T.P., Fujii C.Y., Mason T.M.,
RA Shen M., Ronning C.M., Fraser C.M., Somerville C.R., Venter J.C.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 23-41.
RA Robertson D., Mitchell G.P., Gilroy J.S., Gerrish C., Bolwell G.P.,
RA Slabas A.R.;
RL Submitted (JAN-1997) to the SWISS-PROT data bank.
CC -!- FUNCTION: MAY REMOILIZE PHOSPHATE, PARTICULARLY WHEN CELLS
SENESCE OR WHEN PHOSPHATE BECOMES LIMITING.
CC -!- INDUCTION: BY PHOSPHATE-STARVATION.
CC -!- SIMILARITY: BELONGS THE RNASE T2 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; U05206; AAC48925.1; -;
DR EMBL; AC004138; AAC32917.1; -;
DR HSSP; P08056; IBOL.
DR INTERPRO: IPR001568; -;
DR PFAM; PF00445; ribonuclease_T2; 1.
DR PROSITE; PS00530; RNASE_T2_1; 1.
DR PROSITE; PS00531; RNASE_T2_2; 1.
KW Hydrolyase; Nuclease; Endonuclease; Signal; Multigene family.
FT SIGNAL 1 22
FT CHAIN 23 230 RIBONUCLEASE 1.
FT DISULFID 80 126 BY SIMILARITY.
FT DISULFID 186 221 BY SIMILARITY.
FT ACT_SITE 65 65 BY SIMILARITY.
FT ACT_SITE 119 123 BY SIMILARITY.
FT ACT_SITE 123 123 BY SIMILARITY.
FT CONFLICT 39 41 WPG -> GXP (IN REF. 3).
SQ SEQUENCE 230 AA; 25396 MW; DFD132D39F02505A CRC64;

Query Match 2.3%; Score 7; DB 1; Length 230;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 158 FSASSSS 164
Db |||||
21 FSASSSS 27

RESULT 27
CAS2_PIG
ID CAS2_PIG STANDARD; PRT; 235 AA.
AC P39036;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE ALPHA-S2 CASEIN PRECURSOR.
GN CSN152.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MAMMARY GLAND;
RX MEDLINE=92367960; PubMed=1503276;

RA Alexander L.J., Das Gupta N.A., Beattie C.W.;
RT "The sequence of porcine alpha s2-casein cDNA.";
RL Anim. Genet. 23:365-367(1992).
CC -!- FUNCTION: IMPORTANT ROLE IN THE CAPACITY OF MILK TO TRANSPORT
CALCIUM PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -!- TISSUE SPECIFICITY: MAMMARY GLAND; MILK.
CC -!- SIMILARITY: BELONGS TO THE ALPHA-CASEIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; X54975; CAA38719.1; -;
DR INTERPRO: IPR001588; -;
DR PFAM; PF00363; caseins; 2.
DR PROSITE; PS00306; CASEIN_ALPHA_BETA; 1.
KW Milk; Phosphorylation; Duplication; Signal.
FT SIGNAL 1 15
FT CHAIN 16 235 ALPHA-S2 CASEIN.
FT MOD_RES 23 23 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 24 24 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 25 25 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 32 32 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 70 70 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 71 71 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 72 72 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 73 73 PHOSPHORYLATION (BY SIMILARITY).
FT MOD_RES 74 74 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 235 AA; 27570 MW; C903B760D184C14C CRC64;

Query Match 2.3%; Score 7; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 ASSSSSE 166
Db |||||
69 ASSSSSE 75

RESULT 28
DCOP_PYRHO
ID DCOP_PYRHO STANDARD; PRT; 241 AA.
AC O58462;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROBABLE OROTIDINE 5'-PHOSPHATE DECARBOXYLASE (EC 4.1.1.23) (OMP
DECARBOXYLASE) (OMPDCase).
GN PYRF OR PH0731.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yanamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudo H. Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).
CC -!- CATALYTIC ACTIVITY: OROTIDINE 5'-PHOSPHATE -> UMP + CO(2).
CC -!- PATHWAY: SIXTH AND LAST STEP IN PYRIMIDINE BIOSYNTHESIS.
CC -!- SIMILARITY: BELONGS TO THE OMP DECARBOXYLASE FAMILY.


```
Db 57 CQCPGPG 63
|||||
RESULT 31
Y495_MYCLE STANDARD: PRT: 323 AA.
AC P54579;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYPOTHETICAL 36.1 KDA PROTEIN B2168_F1_37.
GN B2168_F1_37.
OS Mycobacterium leprae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
RN [1]
RA Smith D.R., Robison K.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: STRONG, TO M.TUBERCULOSIS RV0495C.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AJ001202; CA04596.1; -
DR INTERPRO; IPR000488; -
DR INTERPRO; IPR001368; -
DR PFAM; PF00020; TNFR_C6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00050; TNFR_NGFR_2; 2.
DR PROSITE; PS00017; DEATH_DOMAIN; 1.
KW Apoptosis; Receptor; Glycoprotein; Transmembrane; Repeat; Signal.
FT SIGNAL 1 16
FT CHAIN 17 332 FASL RECEPTOR.
FT DOMAIN 17 175 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 176 192 POTENTIAL.
FT DOMAIN 193 332 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 45 164 3 X TNFR-CYS.
FT REPEAT 45 81 TNFR-CYS 1.
FT REPEAT 82 125 TNFR-CYS 2.
FT REPEAT 126 164 TNFR-CYS 3.
FT DOMAIN 227 311 DEATH DOMAIN.
FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 323 AA; 36126 MW; 3F5BA32F82BD9F4C CRC64;

Query Match 2.3%; Score 7; DB 1; Length 323;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 261 LTELIGA 267
|||||
Db 288 LTELIGA 294

RESULT 32
FASA_PIG STANDARD: PRT: 332 AA.
ID FASA_PIG
AC O77736;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE FASL RECEPTOR PRECURSOR (APOPTOSIS-MEDIATING SURFACE ANTIGEN FAS)
DE (APO-1 ANTIGEN) (CD95).
GN TNFRSF6 OR APT1 OR FAS.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN [1]
RP SEQUENCE FROM N.A.
RA Bartling B., Hoffmann J., Holtz J., Schulz R., Heusch G., Darmer D.;
RT "Expression of apoptosis-associated genes in hibernating and stunned
RT myocardium of pig.";
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS FASL. THE
CC ADAPTOR MOLECULE FADD RECRUITS CASPASE-8 TO THE ACTIVATED
CC RECEPTOR. THE RESULTING AGGREGATE CALLED THE DEATH-INDUCING
CC SIGNALING COMPLEX (DISC) PERFORMS CASPASE-8 PROTEOLYTIC
CC ACTIVATION. ACTIVE CASPASE-8 INITIATES THE SUBSEQUENT CASCADE OF
CC CASPASES (ASPARTATE-SPECIFIC CYSTEINE PROTEASES) MEDIATING
CC APOPTOSIS. FAS-MEDIATED APOPTOSIS MAY HAVE A ROLE IN THE
CC INDUCTION OF PERIPHERAL TOLERANCE, IN THE ANTIGEN-STIMULATED
CC SUICIDE OF MATURE T-CELLS, OR BOTH (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.

-!- DOMAIN: CONTAINS A DEATH DOMAIN INVOLVED IN THE BINDING OF FADD,
AND MAYBE TO OTHER CYTOSOLIC ADAPTOR PROTEINS.
-!- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
-!- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AJ001202; CA04596.1; -
DR INTERPRO; IPR000488; -
DR INTERPRO; IPR001368; -
DR PFAM; PF00020; TNFR_C6; 3.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS00050; TNFR_NGFR_2; 2.
DR PROSITE; PS00017; DEATH_DOMAIN; 1.
KW Apoptosis; Receptor; Glycoprotein; Transmembrane; Repeat; Signal.
FT SIGNAL 1 16
FT CHAIN 17 332 FASL RECEPTOR.
FT DOMAIN 17 175 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 176 192 POTENTIAL.
FT DOMAIN 193 332 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 45 164 3 X TNFR-CYS.
FT REPEAT 45 81 TNFR-CYS 1.
FT REPEAT 82 125 TNFR-CYS 2.
FT REPEAT 126 164 TNFR-CYS 3.
FT DOMAIN 227 311 DEATH DOMAIN.
FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 323 AA; 37592 MW; 5B8B03682756BF1B CRC64;

Query Match 2.3%; Score 7; DB 1; Length 332;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 150 CQCPGPG 156
|||||
Db 58 CQCPGPG 64

RESULT 33
SNC2_HUMAN STANDARD: PRT: 334 AA.
ID SNC2_HUMAN
AC Q13487; Q13486;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE SNRPA ACTIVATING PROTEIN COMPLEX 45 KDA SUBUNIT (SNAPC 45 KDA
DE SUBUNIT) (PROXIMAL SEQUENCE ELEMENT-BINDING TRANSCRIPTION FACTOR DELTA
DE SUBUNIT) (PSE-BINDING FACTOR DELTA SUBUNIT) (PTF DELTA SUBUNIT).
GN SNAPC2 OR SNAP45.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-TERATOCARCINOMA;
RX MEDLINE=96210634; PubMed=8633057;
RA Sadowski C.L., Henry R.W., Kobayashi R., Hernandez N.;
RT "The SNAP45 subunit of the small nuclear RNA (snRNA) activating
RT protein complex is required for RNA polymerase II and III snRNA gene
RT transcription and interacts with the TATA box binding protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:4289-4293(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96104548; PubMed=8524284;
RA Yoon J.B., Roeder R.G.;
```

RT "Cloning of two proximal sequence element-binding transcription factor
RT subunits (gamma and delta) that are required for transcription of
RT small nuclear RNA genes by RNA polymerases II and III and interact
RT with the TATA-binding protein.";
CC Mol. Cell. Biol. 16:1-9(1996).
CC
CC -1- FUNCTION: COMPLEX REQUIRED FOR THE TRANSCRIPTION OF BOTH RNA
CC POLYMERASE II AND III SMALL-NUCLEAR RNA GENES. BINDS TO THE
CC PROXIMAL SEQUENCE ELEMENT (PSE), A NON-TATA-BOX BASAL PROMOTER
CC ELEMENT COMMON TO THESE 2 TYPES OF GENES.
CC
CC -1- SUBUNIT: COMPOSED OF 5 SUBUNITS: SNAP190; SNAP50; SNAP45; SNAP43
CC AND SNAP19. SNAP45 INTERACTS STRONGLY WITH TBP.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U44898; AAB06230.1; -;
CC EMBL; U44755; AAC50359.1; -;
CC TRANSFAC; T01681; -;
CC KW Transcription regulation.
CC FT CONFLICT 118 118 L -> V (IN REF. 2).
CC SQ SEQUENCE 334 AA; 35556 MW; 441894AF6E1F4083 CRC64;

Query Match 2.3%; Score 7; DB 1; Length 334;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 273 LVRLQA 279
Db 32 LVRLQA 38
| | | | |

RESULT 34
NOD2_BRASN
ID NOD2_BRASN STANDARD; PRT; 337 AA.
AC P30328;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE NODULATION PROTEIN D II.
GN NODD2.
OS Bradyrhizobium sp. (strain NC92).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bradyrhizobium group; Bradyrhizobium.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96213014; PubMed=8631662;
RA Gillette W.K., Elkan G.H.;
RT "Bradyrhizobium (Arachis) sp. strain NC92 contains two nodD genes
RT involved in the repression of nodA and a nola gene required for the
RT efficient nodulation of host plants.";
RL J. Bacteriol. 178:2757-2766(1996).
CC
CC -1- FUNCTION: NODD REGULATES THE EXPRESSION OF THE NOD ABCPE GENES
CC WHICH ENCODE OTHER NODULATION PROTEINS. NODD IS ALSO A NEGATIVE
CC REGULATOR OF ITS OWN EXPRESSION. BINDS FLAVONOIDS AS INDUCERS.
CC
CC -1- MISCELLANEOUS: THERE ARE AT LEAST TWO NODD GENES IN B.SP STRAIN
CC NC92.
CC
CC -1- SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

DR EMBL; U33192; AAB06563.1; -;
DR INTERPRO; IPR000847; -;
DR INTERPRO; IPR001583; -;
DR PFAM; PF00126; HTH_1; 1.
DR PFAM; PF01046; NODD_C-term; 1.
DR PRINTS; PR00039; HTHLYSR.
DR PROSITE; PS00044; HTH_LYSR_FAMILY; 1.
CC KW Modulation; Transcription regulation; DNA-binding; Activator;
CC KW Repressor; Multigene family; H-T-H MOTIF (POTENTIAL).
FT DNA_BIND 23 42
SQ SEQUENCE 337 AA; 38141 MW; 7645884AB5160177 CRC64;

Query Match 2.3%; Score 7; DB 1; Length 337;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 GERLCA 50
Db 167 GERLCA 173
| | | | |

RESULT 35
HMZ1_DROME
ID HMZ1_DROME STANDARD; PRT; 353 AA.
AC P09089; Q9VI45;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ZERKNUELLT PROTEIN I (ZEN-I).
GN ZEN OR ZENI OR ZI.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88112803; PubMed=2892759;
RA Rushlow C., Doyle H., Hoey T., Levine M.;
RT "Molecular characterization of the zerknullt region of the
RT Antennapedia gene complex in Drosophila.";
RL Genes Dev. 1:1268-1279(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Celniker S.E., Pfeiffer B., Knafels J., Martin C.H., Mayeda C.A.,
RA Palazzolo M.J.;
RT "Complete sequence of the Antennapedia complex of Drosophila.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.A.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler C., Cadenot L.B., Davies P.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos K., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector R., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassaran D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RA "The genome sequence of *Drosophila melanogaster*."
RL Science 287:2185-2195(2000).
CC -!- FUNCTION: REQUIRED FOR THE DIFFERENTIATION OF THE DORSAL-VENTRAL
CC PATTERN, AND DOES NOT APPEAR TO BE INVOLVED IN THE PROCESS OF
CC SEGMENTATION.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X68347; CAA48416.1; -.
DR EMBL: AE001574; AAD19800.1; -.
DR EMBL: AE003674; AAF54087.1; -.
DR PIR: A43697; A43697.
DR HSP: P02833; ISAN.
DR TRANSFAC: T00917; -.
DR FLYBASE: FBgn0004053; zen.
DR INTERPRO: IPR001356; -.
DR PFAM: PF000046; homeobox.1.
DR PRINTS: PR00024; HOMEBOX.
DR PROSITE: PS00027; HOMEBOX_1; 1.
DR PROSITE: PS50071; HOMEBOX_2; 1.
KW Transcription regulation; Homeobox; DNA-binding; Nuclear protein;
KW Developmental protein.
KW DNA_BIND 90 149 HOMEBOX
SQ SEQUENCE 353 AA; 39302 MW; 1FA64031C160CE2B CRC64;

Query Match 2.3%; Score 7; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
| | | | | | |
DB 260 SASSSSS 266

RESULT 36
OTX1_HUMAN
ID OTX1_HUMAN STANDARD; PRT; 354 AA.
AC P32242;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE HOMEBOX PROTEIN OTX1.
GN OTX1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-93327763; PubMed-8101484;

RA Simeone A., Acampora D., Mallamaci A., Stornaiuolo A., D'Apice M.R.,
RA Nigro V., Boncinelli E.;
RT "A vertebrate gene related to orthodenticle contains a homeodomain of
RT the bicoid class and demarcates anterior neuroectoderm in the
RT gastrulating mouse embryo.";
RL EMBO J. 12:2735-2747(1993).
CC -!- FUNCTION: PROBABLY PLAY A ROLE IN THE DEVELOPMENT OF THE BRAIN AND
CC THE SENSE ORGANS. CAN BIND TO THE BCD TARGET SEQUENCE (BTS):
CC 5'-TCTAATCCC-3'.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN.
CC -!- DEVELOPMENTAL STAGE: EMBRYO.
CC -!- SIMILARITY: BELONGS TO THE PAIRED FAMILY OF HOMEBOX PROTEINS.
CC "BICOID" SUBFAMILY.
DR HSP: P06601; 1FJL.
DR TRANSFAC: T02079; -.
DR MIN: 600036; -.
DR INTERPRO: IPR001356; -.
DR INTERPRO: IPR003025; -.
DR INTERPRO: IPR003026; -.
DR PFAM: PF000046; homeobox.1.
DR PRINTS: PR00024; HOMEBOX.
DR PRINTS: PR01255; OTX1HOMEBOX.
DR PROSITE: PR01256; OTX1HOMEBOX.
DR PROSITE: PS00027; HOMEBOX_1; 1.
DR PROSITE: PS50071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein.
KW DNA_BIND 38 97 HOMEBOX.
ET DOMAIN 275 301 HIS-RICH.
SQ SEQUENCE 354 AA; 37327 MW; E32C1E47464BDDAA CRC64;

Query Match 2.3%; Score 7; DB 1; Length 354;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
| | | | | | |
DB 135 SASSSSS 141

RESULT 37
OTX1_MOUSE
ID OTX1_MOUSE STANDARD; PRT; 355 AA.
AC P80205;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE HOMEBOX PROTEIN OTX1.
GN OTX1 OR OTX-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6;
RX MEDLINE-93327763; PubMed-8101484;
RA Simeone A., Acampora D., Mallamaci A., Stornaiuolo A., D'Apice M.R.,
RA Nigro V., Boncinelli E.;
RT "A vertebrate gene related to orthodenticle contains a homeodomain of
RT the bicoid class and demarcates anterior neuroectoderm in the
RT gastrulating mouse embryo.";
RL EMBO J. 12:2735-2747(1993).
CC -!- FUNCTION: PROBABLY PLAY A ROLE IN THE DEVELOPMENT OF THE BRAIN AND
CC THE SENSE ORGANS. CAN BIND TO THE BCD TARGET SEQUENCE (BTS):
CC 5'-TCTAATCCC-3'.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN.
CC -!- DEVELOPMENTAL STAGE: EMBRYO.
CC -!- SIMILARITY: BELONGS TO THE PAIRED FAMILY OF HOMEBOX PROTEINS.
CC "BICOID" SUBFAMILY.
DR HSP: P06601; 1FJL.
DR TRANSFAC: T02079; -.
DR MIN: 600036; -.
DR INTERPRO: IPR001356; -.
DR INTERPRO: IPR003025; -.
DR INTERPRO: IPR003026; -.
DR PFAM: PF000046; homeobox.1.
DR PRINTS: PR00024; HOMEBOX.
DR PRINTS: PR01255; OTX1HOMEBOX.
DR PROSITE: PR01256; OTX1HOMEBOX.
DR PROSITE: PS00027; HOMEBOX_1; 1.
DR PROSITE: PS50071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein.
KW DNA_BIND 38 97 HOMEBOX.
ET DOMAIN 275 301 HIS-RICH.
SQ SEQUENCE 354 AA; 37327 MW; E32C1E47464BDDAA CRC64;

```

CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC -!- TISSUE SPECIFICITY: BRAIN: RESTRICTED REGIONS OF THE DEVELOPING
CC ROSTRAL BRAIN INCLUDING THE PRESUMPTIVE CEREBRAL CORTEX AND
CC OLFACTORY BULBS; EXPRESSED IN THE DEVELOPING OLFACTORY, AURICULAR
CC AND OCULAR SYSTEMS, INCLUDING THE COVERING OF THE OPTIC NERVE.
CC -!- DEVELOPMENTAL STAGE: EMBRYO.
CC -!- SIMILARITY: BELONGS TO THE PAIRED FAMILY OF HOMEOBOX PROTEINS.
CC "BICOID" SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X68883; CAA48754.1; -
CC PIR; S35345; S35345.
CC HSSP; P06601; 1FJL.
CC TRANSFAC; T02080; -
CC MGD; MGI:97450; OTX1.
CC INTERPRO; IPR001356; -
CC INTERPRO; IPR003025; -
CC INTERPRO; IPR003026; -
CC PFAM; PF00046; homeobox; 1.
CC PRINTS; PR00024; HOMEOBOX.
CC PRINTS; PR01255; OTXHOMEOBOX.
CC PRINTS; PR01256; OTXHOMEOBOX.
CC PROSITE; PS00027; HOMEOBOX_1; 1.
CC PROSITE; PS00071; HOMEOBOX_2; 1.
CC HOMEOBOX; DNA-binding; Developmental protein; Nuclear protein.
CC FT DOMAIN 275 302 HIS-RICH.
CC FT BINDING 38 97
CC SEQUENCE 355 AA; 37531 MW; DA5136D9D4FC948F CRC64;

Query Match 2.3%; Score 7; DB 1; Length 355;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
DB 135 SASSSSS 141
|||||

RESULT 38
ID OTX1_RAT STANDARD; PRT; 355 AA.
AC Q63410; Q64203;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HOMEOBOX PROTEIN OTX1.
GN OTX1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=95016961; PubMed=7931541;
RA Frantz G.D., Weimann J.M., Levin M.E., McConnell S.K.;
RT "Otx1 and Otx2 define layers and regions in developing cerebral
RT cortex and cerebellum.";
RL J. Neurosci. 14:5725-5740(1994).
RN [2]
RP SEQUENCE OF 176-323 FROM N.A.
RX MEDLINE=96108898; PubMed=8613727;
RA Robel L., Ding M., James A.J., Lin X., Simeone A., Leckman J.F.,
RA Vaccarino F.M.;
RT "Fibroblast growth factor 2 increases Otx2 expression in precursor
RT cells from mammalian telencephalon.";

```

```

RL J. Neurosci. 15:7879-7891(1995).
CC -!- FUNCTION: PROBABLY PLAY A ROLE IN THE DEVELOPMENT OF THE BRAIN AND
CC THE SENSE ORGANS. CAN BIND TO THE BCD TARGET SEQUENCE (BTS): 5'-
CC TCTAATCC-3' (BY SIMILARITY). MAY PLAY A ROLE IN THE SPECIFICATION
CC OR DIFFERENTIATION OF NEURONS IN THE DEEP LAYERS OF THE CEREBRAL
CC CORTEX, AND ALSO IN CEREBELLAR REGIONALIZATION DURING EARLY
CC DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC -!- TISSUE SPECIFICITY: BRAIN: RESTRICTED REGIONS OF THE DEVELOPING
CC ROSTRAL BRAIN INCLUDING THE PRESUMPTIVE CEREBRAL CORTEX AND
CC OLFACTORY BULBS; EXPRESSED IN THE DEVELOPING OLFACTORY, AURICULAR
CC AND OCULAR SYSTEMS, INCLUDING THE COVERING OF THE OPTIC NERVE (BY
CC SIMILARITY).
CC -!- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT THE FOREBRAIN AND
CC MIDBRAIN DURING DEVELOPMENT, AND IN ADDITION IS ALSO SEEN IN
CC DISCRETE SPATIAL AND TEMPORAL DOMAINS IN THE DEVELOPING CEREBRAL
CC CORTEX AND CEREBELLUM. CONFINED TO A SUBPOPULATION OF NEURONS IN
CC LAYERS 5 AND 6 WITHIN THE ADULT CEREBRAL CORTEX AND DURING
CC DEVELOPMENT EXPRESSION IS HIGH IN THE PROGENITORS OF THESE DEEP-
CC LAYER CELLS. EXPRESSED IN THE DEVELOPING CEREBELLUM IN SPATIALLY
CC RESTRICTED REGIONS OF THE EXTERNAL GRANULAR LAYER.
CC -!- SIMILARITY: BELONGS TO THE PAIRED FAMILY OF HOMEOBOX PROTEINS.
CC "BICOID" SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L32602; AAA53557.1; -
CC EMBL; S81924; -; NOT_ANNOTATED_CDS.
CC HSSP; P06601; 1FJL.
CC INTERPRO; IPR001356; -
CC INTERPRO; IPR003025; -
CC INTERPRO; IPR003026; -
CC PFAM; PF00046; homeobox; 1.
CC PRINTS; PR00024; HOMEOBOX.
CC PRINTS; PR01255; OTXHOMEOBOX.
CC PRINTS; PR01256; OTXHOMEOBOX.
CC PROSITE; PS00027; HOMEOBOX_1; 1.
CC PROSITE; PS00071; HOMEOBOX_2; 1.
CC HOMEOBOX; DNA-binding; Developmental protein; Nuclear protein.
CC FT DNA_BIND 38 97 HOMEOBOX.
CC FT DOMAIN 275 302 HIS-RICH.
CC FT CONFLICT 219 219 T -> A (IN REF. 2).
CC FT CONFLICT 229 229 R -> G (IN REF. 2).
CC SEQUENCE 355 AA; 37602 MW; C875871723D0B876 CRC64;

Query Match 2.3%; Score 7; DB 1; Length 355;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
DB 135 SASSSSS 141
|||||

RESULT 39
ID CKRA_HUMAN STANDARD; PRT; 362 AA.
AC P46092;
DT 01-NOV-1995 (Rel. 32, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE C-C-CHEMOKINE RECEPTOR TYPE 10 (C-C CKR-10) (CCR-10).
GN CCR10 OR GPR2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

RN SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=20191997; PubMed=10725696;
RA Jarmin D.I., Rits M., Bota D., Gerard N.P., Graham G.J.,
RA Clark-Lewis I., Gerard C.;
RT "Cutting edge: Identification of the orphan receptor G-protein-coupled
RT receptor 2 as CCR10, a specific receptor for the chemokine ESKine.";
RL J. Immunol. 164:3460-3464(2000).
RN [2]
RN SEQUENCE OF 9-362 FROM N.A.
RX MEDLINE=95154831; PubMed=7851889;
RA Marchese A., Docherty J.M., Nguyen T., Heiber M., Cheng R.,
RA Heng H.H.Q., Tsui L.-C., Shi X., George S.R., O'Dowd B.F.;
RT "Cloning of human genes encoding novel G protein-coupled receptors.";
RL Genomics 23:609-618(1994).
RN [3]
RN CHARACTERIZATION.
RX MEDLINE=20191998; PubMed=10725697;
RA Honey B., Wang W., Soto H., Buchanan M.E., Wiesenborn A., Catron D.,
RA Muller A., McClanahan T.K., Dieu-Nosjean M.C., Orozco R., Ruzicka T.,
RA Lehmann P., Oldham E., Zlotnik A.;
RT "Cutting edge: the orphan chemokine receptor G protein-coupled
RT receptor-2 (GPR-2, CCR10) binds the skin-associated chemokine CCL27
RT (CTACK/ALP/ILC).";
RL J. Immunol. 164:3465-3470(2000).
CC -|- FUNCTION: RECEPTOR FOR CHEMOKINES SCYA27 AND SCYA28. SUBSEQUENTLY
CC TRANSDUCES A SIGNAL BY INCREASING THE INTRACELLULAR CALCIUM IONS
CC LEVEL.
CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -|- TISSUE SPECIFICITY: EXPRESSED BY MELANOCYTES, DERMAL FIBROBLASTS,
CC AND DERMAL MICROVASCULAR ENDOTHELIAL CELLS. ALSO DETECTED IN T
CC CELLS AS WELL AS IN SKIN-DERIVED LANGERHANS CELLS.
CC -|- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF215981; AAF63709.1; -
DR EMBL; U13667; AAA64593.1; -
DR MIM; 600240; -
DR INTERPRO; IPR000276; -
DR PFAM; PF00001; 7cml_1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECF_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECF_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 52
FT TRANSMEM 53 68
FT DOMAIN 69 78
FT TRANSMEM 79 99
FT DOMAIN 100 114
FT TRANSMEM 115 136
FT DOMAIN 137 159
FT TRANSMEM 160 179
FT DOMAIN 180 203
FT TRANSMEM 204 225
FT DOMAIN 226 247
FT TRANSMEM 248 269
FT DOMAIN 270 290
FT TRANSMEM 291 313
FT DOMAIN 314 362
FT DISULFID 113 191
FT CONFLICT 51 51
V -> L (IN REF. 2).
SQ SEQUENCE 362 AA; 38399 MW; 69DF12B639AE99A CRC64;

Query Match 2.3%; Score 7; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 16 LALPALL 22
Db 172 LALPALL 178
RESULT 40
YK03_CAEEL STANDARD; PRT; 401 AA.
AC P34291;
DT 01-FEB-1994 (Rel. 28, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYPOTHETICAL 43.5 KDA PROTEIN C05B5.3 IN CHROMOSOME III.
GN C05B5.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Mortimore B.J.;
RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RC STRAIN-BRISTOL N2;
RA Durbin R.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z32679; CAA83596.1; -
DR WORMPEP; C05B5.3; C017369.
KW Hypothetical protein.
SQ SEQUENCE 401 AA; 43472 MW; C99DF259EF6C8B55 CRC64;
Query Match 2.3%; Score 7; DB 1; Length 401;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 182 NYPGSSS 188
Db 99 NYPGSSS 105
RESULT 41
HEM1_AQUAE STANDARD; PRT; 406 AA.
ID HEM1_AQUAE
AC O67314;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GLUTAMYL-TRNA REDUCTASE (EC 1.2.1.-) (GLUTR).
GN HEMA OR AQ.1279.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aufay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RT


```
RL Nature 392:353-358(1998).
CC -1- CATALYTIC ACTIVITY: GLUTAMYL-TRNA( GLU ) + NADPH = GLUTAMATE-1-
CC SEMIALDEHYDE + NADP(+) + TRNA( GLU ).
CC -1- PATHWAY: FIRST STEP IN PORPHYRIN BIOSYNTHESIS BY THE C5 PATHWAY.
CC -1- SIMILARITY: BELONGS TO THE GLUTAMYL-TRNA REDUCTASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000732; AAC07274.1; -.
CC HSP; Q42843; J829.
CC INTERPRO: IPR000343; -.
CC PFAM: PF00745; GlutR; 1.
CC PROSITE: PS00747; GLUTR; 1.
CC PORPHYRIN BIOSYNTHESIS: Oxidoreductase; NADP.
CC SEQUENCE 406 AA; 46648 MW; 77F3ECC488FC4394 CRC64;
SQ -----

Query Match 2.3%; Score 7; DB 1; Length 406;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 289 ERSVRER 295
Db 260 ERSVRER 266
|||||||

RESULT 42
SR40_YEAST STANDARD; PRT; 406 AA.
ID SR40_YEAST
AC P32583;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE SUPPRESSOR PROTEIN SRP40.
GN SRP40 OR YKR092C OR YKR412A.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC Saccharomycetaceae; Saccharomycetes.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 28583 / FL100;
RA Lalo D., Carles C., Sentenac A., Thuriaux P.;
RL Submitted (MAY-1993) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94205265; PubMed=8154186;
RA Bou G., Esteban P.F., Baladron V., Gonzalez G.A., Cantalejo J.G.,
RA Remacha M., Jimenez A., del Rey F., Ballesta J.P.G., Revuelta J.L.;
RT "The complete sequence of a 15,820 bp segment of Saccharomyces
RT cerevisiae chromosome XI contains the UBI2 and MPL1 genes and three
RT new open reading frames."
RL Yeast 9:1349-1354(1993).
CC -1- FUNCTION: NOT KNOWN. WEAK SUPPRESSOR OF A MUTANT OF THE
CC SUBUNIT AC40 OF DNA DEPENDANT RNA POLYMERASE I AND III.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L11275; AAA35091.1; -.
CC EMBL; X73541; CAA51946.1; -.
CC EMBL; Z28317; CAA82171.1; -.
CC PIR; S38170; S38170.
```

```
DR SGD; S0001800; SRP40.
FT DOMAIN 25 314 ASP/SER-RICH.
FT CONFLICT 400 400 G -> N (IN REF. 1).
FT SEQUENCE 406 AA; 41015 MW; 8EA007695AF4BA1D CRC64;
SQ -----

Query Match 2.3%; Score 7; DB 1; Length 406;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
Db 308 SASSSSS 314
|||||||

RESULT 43
HID_DROME STANDARD; PRT; 410 AA.
ID HID_DROME
AC Q24106;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE HEAD INVOLUTION DEFECTIVE PROTEIN (WRINKLED PROTEIN).
GN W OR HID.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CANTON-S; TISSUE=EYE IMAGINAL DISK;
RX MEDLINE=95347579; PubMed=7622034;
RA Grether M.E., Abrams J.M., Agapite J., White K., Steller H.;
RT "The head involution defective gene of Drosophila melanogaster
RT functions in programmed cell death."
RL Genes Dev. 9:1694-1708(1995).
CC -1- FUNCTION: REQUIRED FOR INDUCTION OF APOPTOSIS. HID MUTANTS CONTAIN
CC EXTRA CELLS IN THE HEAD DUE TO DECREASED LEVELS OF CELL DEATH
CC AND SHOW A PRONOUNCED DEFECT IN THE MORPHOGENETIC MOVEMENTS OF
CC HEAD INVOLUTION. ECTOPIC EXPRESSION IN THE RETINA RESULTS IN
CC COMPLETE EYE ABLATION. SEEMS TO ACT GENETICALLY UPSTREAM OF
CC BACULOVIRAL ANTI-APOPTOTIC P35.
CC -1- DEVELOPMENTAL STAGE: EXPRESSION CORRESPONDS APPROXIMATELY TO THE
CC PATTERN OF PROGRAMMED CELL DEATH IN THE EMBRYO, PARTICULARLY IN
CC THE HEAD.
CC -1- SIMILARITY: LIMITED AT THE N-TERMINAL REGION, TO GRIM AND RPR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U31226; AAA79985.1; -.
CC FLYBASE; FBgn0003997; W.
KW Apoptosis; Developmental protein; Polymorphism.
FT DOMAIN 17 20 POLY-SER.
FT DOMAIN 50 62 POLY-SER.
FT DOMAIN 237 240 POLY-SER.
FT DOMAIN 332 340 POLY-SER.
FT VARIANT 171 171 P -> S (IN A22 ALLELE).
FT VARIANT 261 261 S -> L (IN A206 ALLELE).
FT SEQUENCE 410 AA; 43866 MW; 63EBF910179E24E2 CRC64;
SQ -----

Query Match 2.3%; Score 7; DB 1; Length 410;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
|||||||
```



```
Db 53 SASSSS 59
RESULT 44
CLPX_AQUAE STANDARD; PRT; 412 AA.
AC 067356;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ATP-DEPENDENT CLP PROTEASE ATP-BINDING SUBUNIT CLPX.
GN CLPX OR AQ_1337.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus";
RL Nature 392:353-358(1998).
CC -!- FUNCTION: ATP-DEPENDENT SPECIFICITY COMPONENT OF THE CLP PROTEASE.
CC IT DIRECTS THE PROTEASE TO SPECIFIC SUBSTRATES. CAN PERFORM
CC CHAPERONE FUNCTIONS IN THE ABSENCE OF CLPP (BY SIMILARITY).
CC -!- SUBUNIT: HETERODIMER OF CLPP AND CLPX (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CLPX CHAPERONE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000735; AAC07316.1; -
DR INTERPRO; IPR001939; -
DR PFAM; PF00004; AAA; 1.
KW Chaperone; ATP-binding.
FT ZN_FING 10 36 C4-TYPE.
FT NP_BIND 117 124 ATP (POTENTIAL).
SQ SEQUENCE 412 AA; 46140 MW; FFCFAF451A930B69 CRC64;

Query Match 2.3%; Score 7; DB 1; Length 412;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 273 LVRLLOA 279
Db 160 LVRLLOA 166
|||||||

RESULT 45
HXD3_MOUSE STANDARD; PRT; 417 AA.
AC P09027;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HMOBOX PROTEIN HOX-D3 (HOX-4.1) (MH-19).
GN HOXD3 OR HOXD-3 OR HOX-4.1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94368862; PubMed=7916214;
RA Brown W.M., Zhou L., Taylor G.R.;
RT "The nucleotide sequence of the murine Hox-D3 (Hox-4.1) gene reveals
extensive identity with the human protein.";
Biochim. Biophys. Acta 1219:219-222(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95225976; PubMed=7710686;
RA Bedford M., Orr-Urtreger A., Arman E., Lonai P.;
RT "Analysis of the Hoxd-3 gene: structure and localization of its sense
RT and natural antisense transcripts.";
RL DNA Cell Biol. 14:295-304(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR SWISS;
RA Tan D., Shao X., Pu L., Guo V., Nirenberg M.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 167-267 FROM N.A.
RC STRAIN=ICR SWISS; TISSUE=LIVER;
RA Tan D.P., Shao X., Nirenberg M.;
RL Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 169-290 FROM N.A.
RX MEDLINE=88054465; PubMed=2890503;
RA Lonai P., Arman E., Czosnek H., Ruddle F.H., Blatt C.;
RT "New murine homeoboxes: structure, chromosomal assignment, and
RT differential expression in adult erythropoiesis.";
RL DNA 6:409-418(1987).
RN [6]
RP SEQUENCE OF 166-322 FROM N.A.
RA Lonai P.;
RL Submitted (DEC-1987) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF
CC A DEVELOPMENTAL REGULATORY SYSTEM THAT PROVIDES CELLS WITH
CC SPECIFIC POSITIONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- SIMILARITY: BELONGS TO THE ANTP FAMILY OF HMOBOX PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U03485; AAB60683.1; -
DR EMBL; U03496; AAB60683.1; JOINED.
DR EMBL; X73573; CAAS1975.1; -
DR EMBL; L24363; AAA37855.1; -
DR EMBL; M18169; AAA37846.1; -
DR EMBL; U56400; AAC52779.1; -
DR EMBL; U56401; AAC52780.1; -
DR PIR; D29585; D29585.
DR MGI; MGI:96207; HOXD3.
DR INTERPRO; IPR001356; -
DR INTERPRO; IPR001827; -
DR PFAM; PF00046; homeobox; 1.
DR PRINTS; PR00024; HMOBOX.
DR PRINTS; PR00025; ANTENNAPEDIA.
DR PROSITE; PS00027; HMOBOX_1; 1.
DR PROSITE; PS00071; HMOBOX_2; 1.
DR PROSITE; PS00032; ANTENNAPEDIA; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein;
Transcription regulation.
FT DOMAIN 145 150 ANTP-TYPE HEXAPEPTIDE.
FT DNA_BIND 179 238 HMOBOX.
FT CONFLICT 171 171 D -> N (IN REF. 5 AND 6).
FT CONFLICT 176 176 G -> D (IN REF. 5 AND 6).
FT CONFLICT 185 185 A -> T (IN REF. 5 AND 6).
FT CONFLICT 273 273 MISSING (IN REF. 2).
FT CONFLICT 273 288 PVPGIAYDAPSPAPA -> RCPAXLRRLAARF (IN
REF. 5 AND 6).
FT CONFLICT 298 301 AAYT -> PPTR (IN REF. 6).
FT CONFLICT 298 417 AAYTAPLSSCLPQOKRYPAPEFEPHPMASNGGFASANLQ
```

FT SPVYGGNEVDMAPTSGPVFNGLHSHSPSSVDYSCAAQ
 FT IFCNHHGCDPHPTVTDLSAHSSOGRLPEAPKLTHL ->
 FT GLHGAQQLPASAEAVPGARVRASPHGEORRASPAPTCAA
 FT RCTWATSTFWPRPRGSSIWATSKTRLRPAWTTVAPRKS
 FT LATITTDPAITLPTQISRLTLRLRACPRPN (IN
 FT REF. 2).
 SQ SEQUENCE 417 AA; 44142 MW; 4B6BFEBF1406379 CRC64;

Query Match 2.3%; Score 7; DB 1; Length 417;
 Best Local Similarity 100.0%; Pred. No. 58;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
 DB 133 SASSSSS 139
 |||||

RESULT 46
 PEL_LILLO STANDARD; PRT; 434 AA.
 AC P40973;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE PECTATE LYASE PRECURSOR (EC 4.2.2.2).
 OS Lilium longiflorum (Trumpet lily).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Liliales; Liliaceae; Lilium.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NELLIE WHITE; TISSUE=POLLEN;
 RA Kim S.R., Finkel D.J., An G.;
 RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: ELIMINATIVE CLEAVAGE OF PECTATE TO GIVE
 CC OLIGOSACCHARIDES WITH 4-DEOXY-ALPHA-D-MANN-4-ENURONOSYL GROUPS
 CC AT THEIR NON-REDUCING ENDS.
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z17328; CAA78976.1; -
 CC EMBL; L18911; AAA33398.1; -
 CC PIR; S29612; S29612.
 CC INTERPRO: IPR002022; -
 CC PFAM: PF00544; pec_lyase; 1.
 CC PRINTS: PR00807; AMBALLERGEN.
 KW Lyase; Signal.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 434 PECTATE LYASE.
 FT ACT_SITE 312 312 POTENTIAL.
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 434 AA; 48457 MW; C1F3E30AD2BBD064 CRC64;

Query Match 2.3%; Score 7; DB 1; Length 434;
 Best Local Similarity 100.0%; Pred. No. 60;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 PGTFVQR 60
 DB 412 PGTFVQR 418
 |||||

RESULT 47
 GSA_HORVU STANDARD; PRT; 469 AA.

ID GSA_HORVU STANDARD; PRT; 469 AA.
 AC P18492;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE GLUTAMATE-1-SEMIALDEHYDE 2,1-AMINOMUTASE PRECURSOR (EC 5.4.3.8) (GSA)
 DE (GLUTAMATE-1-SEMIALDEHYDE AMINOTRANSFERASE) (GSA-AT).
 GN GSA.
 OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Hordeum.
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=CV. BONUS; TISSUE=SEEDLING;
 RX MEDLINE=90272675; PubMed=2349227;
 RA Grimm B.;
 RT "Primary structure of a key enzyme in plant tetrapyrrole synthesis:
 RT glutamate 1-semialdehyde aminotransferase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:4169-4173(1990).
 RN [2]
 RP PARTIAL SEQUENCE.
 RX MEDLINE=89374545; PubMed=2505791;
 RA Grimm B., Bull A., Welinder K.G., Gough S.P., Kannangara C.G.;
 RT "Purification and partial amino acid sequence of the glutamate 1-
 RT semialdehyde aminotransferase of barley and synechococcus.";
 RL Carlsberg Res. Commun. 54:67-79(1989).
 CC -1- CATALYTIC ACTIVITY: (S)-4-AMINO-5-OXOPENTANOATE -
 CC 5-AMINOLEVULINATE.
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
 CC -1- PATHWAY: SECOND STEP IN PORPHYRIN BIOSYNTHESIS BY THE C5 PATHWAY.
 CC INVOLVED IN CHLOROPHYLL BIOSYNTHESIS.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.
 CC -1- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT
 CC AMINOTRANSFERASES.
 CC -1- CAUTION: THE RESIDUE IN POSITION 309 IS GLN IN REF.1, BUT IS
 CC SAID TO BE LYS IN FURTHER PAPERS BY THE SAME GROUP.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M31545; AAB59330.1; -
 CC PIR; A35789; A35789.
 CC HSP; P24630; 3GSA.
 CC INTERPRO: IPR000954; -
 CC PFAM: PF00202; aminotran_3; 1.
 CC PROSITE: PS00600; AA_TRANSFER_CLASS_3; 1.
 CC PORPHYRIN BIOSYNTHESIS; Chlorophyll biosynthesis; Isomerase;
 KW Pyridoxal phosphate; Transit peptide; Chloroplast.
 FT TRANSIT 1 34 CHLOROPLAST.
 FT CHAIN 35 469 GLUTAMATE-1-SEMIALDEHYDE 2,1-AMINOMUTASE.
 FT BINDING 309 309 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 FT CONFLICT 309 309 K -> Q (IN REF. 1; SEE COMMENT).
 SQ SEQUENCE 469 AA; 49494 MW; 78BF03C14A3C1448 CRC64;

Query Match 2.3%; Score 7; DB 1; Length 469;
 Best Local Similarity 100.0%; Pred. No. 65;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 268 QDGALLV 274
 DB 273 QDGALLV 279
 |||||

RESULT 48
 YBS8_YEAST STANDARD; PRT; 469 AA.
 ID YBS8_YEAST

P38248;
01-OCT-1994 (Rel. 30, Created)
01-FEB-1996 (Rel. 33, Last sequence update)
15-JUL-1998 (Rel. 36, Last annotation update)
DE HYPOTHETICAL 48.3 KDA PROTEIN IN HSP26-IF32 INTERGENIC REGION.
GN YBR078W OR YBR0727.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC Saccharomycetaceae; Saccharomycetes.
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=S288C;
RX MEDLINE=95076715; PubMed=7985423;
RA van der Aart Q.J.M., Barthe C., Dolignon F., Aigle M., Crouzet M.,
RA Steensma H.Y.;
RT "Sequence analysis of a 31 kb DNA fragment from the right arm of
RT Saccharomyces cerevisiae chromosome II.";
RL Yeast 10:959-964(1994).
[2]
RN REVISIONS.
RP STRAIN=S288C;
RC STRAIN=S288C;
RA van der Aart Q.J.M.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO YEAST SPORULATION-SPECIFIC PROTEIN 2 (SPS2).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X76294; CAA53935.1; -
DR EMBL: Z35947; CAA85022.1; -
DR PIR: S45474; S45474.
DR SGD: S0000282; ECM33.
KW Hypothetical protein.
FT CONFLICT 21 21 MISSING (IN CAA85022).
SQ SEQUENCE 469 AA; 48307 MW; D6E6F0309EB586CA CRC64;

Query Match 2.3%; Score 7; DB 1; Length 469;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
Db 394 SASSSSS 400
|||||||
RESULT 49
KLF4_HUMAN STANDARD; PRT; 470 AA.
AC O43474; P78338;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE KRUPPEL-LIKE FACTOR 4 (EPITHELIAL ZINC-FINGER PROTEIN EZF).
GN KLF4 OR EZF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=98086278; PubMed=9422764;
RA Yet S.-F., McA'Nulty M.M., Folta S.C., Yen H.-W., Yoshizumi M.,
RA Hsieh C.-M., Layne M.D., Chin M.T., Wang H., Perrella M.A., Jain M.K.,
RA Lee M.-E.;
RT "Human EZF, a Kruppel-like zinc finger protein, is expressed in
RT vascular endothelial cells and contains transcriptional activation
RT and repression domains.";
RL J. Biol. Chem. 273:1026-1031(1998).

[2]
RN SEQUENCE FROM N.A.
RP TISSUE=PLACENTA;
RA Garrett-Sinha L.A., de Crombrughe B.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAY ACT AS A TRANSCRIPTIONAL ACTIVATOR. BINDS THE CACCC
CC CORE SEQUENCE. MAY BE INVOLVED IN THE DIFFERENTIATION OF
CC EPITHELIAL CELLS AND MAY ALSO FUNCTION IN THE DEVELOPMENT OF THE
CC SKELETON AND KIDNEY.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF022184; AAC03462.1; -
DR EMBL: U70663; AAB48399.1; -
DR MIN: 602253; -
DR INTERPRO: IPR000822; -
DR PFAM: PF000096; zf-C2H2; 3.
DR PRINTS: PR000048; ZINCFINGER.
DR PROSITE: PS00028; ZINC_FINGER_C2H2; 3.
KW Transcription regulation; Activator; Zinc-finger; Metal-binding;
KW DNA-binding; Nuclear protein; Repeat.
FT DOMAIN 387 469 ZINC-FINGER.
FT ZN_FING 387 411 C2H2-TYPE.
FT ZN_FING 417 441 C2H2-TYPE.
FT ZN_FING 447 469 C2H2-TYPE.
FT CONFLICT 51 52 AG -> GR (IN REF. 2).
FT CONFLICT 68 68 A -> G (IN REF. 2).
FT CONFLICT 242 242 T -> S (IN REF. 2).
FT CONFLICT 295 295 D -> N (IN REF. 2).
FT CONFLICT 306 306 T -> S (IN REF. 2).
FT CONFLICT 312 312 L -> F (IN REF. 2).
FT CONFLICT 320 320 D -> E (IN REF. 2).
SQ SEQUENCE 470 AA; 50021 MW; BF7CCE8400AFE71 CRC64;

Query Match 2.3%; Score 7; DB 1; Length 470;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
Db 120 SASSSSS 126
|||||||
RESULT 50
KLF4_MOUSE STANDARD; PRT; 474 AA.
AC O60793; P70421;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE KRUPPEL-LIKE FACTOR 4 (GUT ENRICHED KRUPPEL-LIKE FACTOR) (EPITHELIAL
DE ZINC-FINGER PROTEIN EZF).
GN KLF4 OR GLF OR EZF OR ZIE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=96355308; PubMed=8702718;
RA Shields J.W., Christy R.J., Yang V.W.;
RT "Identification and characterization of a gene encoding a
RT gut-enriched Kruppel-like factor expressed during growth arrest.";
RL J. Biol. Chem. 271:20009-20017(1996).

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-B6D2F1; TISSUE-EMBRYONIC FIBROBLAST;
RX MEDLINE-97094913; PubMed-8940147;
RA Garrett-Sinha L.A.; Ebertsperger H., Seldin M.F., de Crombrughe B.;
RT "A gene for a novel zinc-finger protein expressed in differentiated
RT epithelial cells and transiently in certain mesenchymal cells.";
RL J. Biol. Chem. 271:31384-31390(1996).
CC -!- FUNCTION: MAY ACT AS A TRANSCRIPTIONAL ACTIVATOR. BINDS THE CACCC
CC CORE SEQUENCE. MAY BE INVOLVED IN THE DIFFERENTIATION OF
CC EPITHELIAL CELLS AND MAY ALSO FUNCTION IN THE DEVELOPMENT OF THE
CC SKELETON AND KIDNEY.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN THE COLON. LOWER LEVELS
CC IN TESTIS, LUNG AND SMALL INTESTINE.
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U20344; AAC04892.1; ALT_INIT.
DR EMBL; U70662; AAC52939.1; ALT_INIT.
DR MGD; MGI:1342287; KLF4.
DR INTERPRO; IPR000822; .
DR PFAM; PF00096; zf-C2H2; 3.
DR PRINTS; PR00048; ZINC_FINGER.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 3.
KW Transcription regulation; Activator; Zinc-finger; Metal-binding;
KW DNA-binding; Nuclear protein; Repeat.
FT DOMAIN 393 473 ZINC FINGERS.
FT ZN_FING 393 415 C2H2-TYPE.
FT ZN_FING 423 445 C2H2-TYPE.
FT ZN_FING 453 473 C2H2-TYPE.
SQ SEQUENCE 474 AA; 50882 MW; 675F4CDDC89015BA CRC64;

Query Match 2.3%; Score 7; DB 1; Length 474;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
Db 117 SASSSSS 123

Search completed: January 30, 2001, 17:10:43
Job time: 227 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 17:06:21 ; Search time 42.13 Seconds
(without alignments)
834.616 Million cell updates/sec

Title: US-09-518-931-2
Perfect score: 300
Sequence: 1 MRALGPGSLCLVLALPA.....RVAMPGLSVRERFLPVH 300

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 374700 seqs, 117207915 residues

Word size : 6

Total number of hits satisfying chosen parameters: 3268

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

SPTREMBL_15:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_unclassified:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	300	100.0	300	4	Q95407 homo sapien
2	9	3.0	430	6	Q9N092
3	9	3.0	561	10	Q9SH82
4	8	2.7	54	12	Q9SH82 arabidopsis
5	8	2.7	54	12	Q9SH88
6	8	2.7	54	12	Q9SH88 budgerigar
7	8	2.7	54	12	Q9SH89 budgerigar
8	8	2.7	54	12	Q9SH89 budgerigar
9	8	2.7	54	12	Q9SH89 budgerigar
10	8	2.7	54	12	Q9SH89 budgerigar
11	8	2.7	54	12	Q9SH89 budgerigar
12	8	2.7	54	12	Q9SH89 budgerigar
13	8	2.7	54	12	Q9SH89 budgerigar
14	8	2.7	54	12	Q9SH89 budgerigar
15	8	2.7	54	12	Q9SH89 budgerigar
16	8	2.7	54	12	Q9SH89 budgerigar
17	8	2.7	54	12	Q9SH89 budgerigar
18	8	2.7	54	12	Q9SH89 budgerigar
19	8	2.7	54	12	Q9SH89 budgerigar

20	8	2.7	54	12	089904	089904 budgerigar
21	8	2.7	129	6	Q9TTA6	Q9TTA6 sus scrofa
22	8	2.7	145	12	Q9WG04	Q9WG04 budgerigar
23	8	2.7	145	12	Q9IBL0	Q9IBL0 budgerigar
24	8	2.7	181	10	065446	065446 arabidopsis
25	8	2.7	206	2	Q9X6H9	Q9X6H9 streptococ
26	8	2.7	241	10	049719	049719 arabidopsis
27	8	2.7	341	5	Q9XXL8	Q9XXL8 caenorhabdi
28	8	2.7	440	2	Q9JP96	Q9JP96 rhodocyclu
29	8	2.7	477	4	Q9I577	Q9I577 homo sapien
30	8	2.7	531	10	042582	042582 arabidopsis
31	8	2.7	531	10	Q9ST23	Q9ST23 arabidopsis
32	8	2.7	587	12	Q9WG03	Q9WG03 budgerigar
33	8	2.7	599	12	Q9IBL1	Q9IBL1 budgerigar
34	8	2.7	599	12	Q9IBK8	Q9IBK8 budgerigar
35	8	2.7	644	13	Q90974	Q90974 gallus gall
36	8	2.7	781	3	Q9P5G4	Q9P5G4 neurospora
37	8	2.7	1013	2	Q9L953	Q9L953 streptococ
38	8	2.7	4472	2	Q93954	Q93954 streptomyc
39	7	2.3	42	3	Q9P7A7	Q9P7A7 schizosacch
40	7	2.3	56	6	Q9TR13	Q9TR13 sus scrofa
41	7	2.3	82	5	076201	076201 phoneutria
42	7	2.3	82	6	Q28411	Q28411 felis silve
43	7	2.3	87	10	Q9S9T1	Q9S9T1 arabidopsis
44	7	2.3	90	5	Q9VQP7	Q9VQP7 drosophila
45	7	2.3	98	5	Q9NFM0	Q9NFM0 plasmidium
46	7	2.3	99	8	Q9NFM0	Q9NFM0 scenedesmus
47	7	2.3	100	8	Q9MD21	Q9MD21 scenedesmus
48	7	2.3	106	2	Q9PEH3	Q9PEH3 xylella fas
49	7	2.3	112	1	Q9YAE2	Q9YAE2 aeropyrum p
50	7	2.3	115	6	Q29173	Q29173 sus scrofa
51	7	2.3	120	2	Q9XDF1	Q9XDF1 streptomyc
52	7	2.3	123	3	Q9UVF0	Q9UVF0 yarrowia li
53	7	2.3	140	1	Q9YD54	Q9YD54 aeropyrum p
54	7	2.3	140	5	Q9XWP9	Q9XWP9 caenorhabdi
55	7	2.3	149	2	050557	050557 actinobacil
56	7	2.3	150	2	050555	050555 actinobacil
57	7	2.3	150	2	Q43893	Q43893 actinobacil
58	7	2.3	150	2	Q85514	Q85514 actinobacil
59	7	2.3	164	10	Q9STV7	Q9STV7 arabidopsis
60	7	2.3	166	5	Q22678	Q22678 caenorhabdi
61	7	2.3	168	12	Q90342	Q90342 rhesus cyto
62	7	2.3	169	5	Q9VRR7	Q9VRR7 drosophila
63	7	2.3	172	10	Q9LEC7	Q9LEC7 solanum tub
64	7	2.3	173	2	Q9RJG5	Q9RJG5 streptomyc
65	7	2.3	179	12	055569	055569 leucania se
66	7	2.3	182	3	013520	013520 saccharomyc
67	7	2.3	182	5	Q21133	Q21133 caenorhabdi
68	7	2.3	189	2	Q9XCL0	Q9XCL0 streptococ
69	7	2.3	194	2	Q9X6U3	Q9X6U3 streptococ
70	7	2.3	196	10	Q9ZPT2	Q9ZPT2 arabidopsis
71	7	2.3	198	5	Q9Y0A7	Q9Y0A7 drosophila
72	7	2.3	198	5	Q9VJ76	Q9VJ76 drosophila
73	7	2.3	199	2	052946	052946 bacillus su
74	7	2.3	199	5	061173	061173 entodinium
75	7	2.3	199	9	Q9MBI3	Q9MBI3 bacterioph
76	7	2.3	201	2	Q31854	Q31854 bacillus su
77	7	2.3	208	1	Q90235	Q90235 pyrococcus
78	7	2.3	208	2	Q9X6I2	Q9X6I2 streptococ
79	7	2.3	209	10	Q9ZNX3	Q9ZNX3 petunia hyb
80	7	2.3	211	3	Q12549	Q12549 aspergillus
81	7	2.3	215	11	061685	061685 mus musculu
82	7	2.3	224	5	Q9W0T8	Q9W0T8 drosophila
83	7	2.3	225	10	082315	082315 arabidopsis
84	7	2.3	226	12	P88823	P88823 salmirine
85	7	2.3	229	11	Q9JM93	Q9JM93 mus musculu
86	7	2.3	230	5	045766	045766 caenorhabdi
87	7	2.3	233	5	Q9XYW8	Q9XYW8 entodinium
88	7	2.3	236	10	004328	004328 arabidopsis
89	7	2.3	237	10	Q9SIA9	Q9SIA9 arabidopsis
90	7	2.3	239	2	Q9L7T6	Q9L7T6 rhodospiril
91	7	2.3	242	3	P89498	P89498 saccharomyc
92	7	2.3	243	4	Q9P2R2	Q9P2R2 homo sapien

93	7	2.3	243	6	Q9N2G3	Q9n2g3 pan troglod	166	7	2.3	410	5	Q9VVP1	Q9vvp1 drosophila
94	7	2.3	243	6	Q9N2G2	Q9n2g2 gorilla gor	167	7	2.3	413	12	Q9INE3	Q9ine3 white spot
95	7	2.3	243	6	Q9N2G1	Q9n2g1 pongo pygma	168	7	2.3	414	4	O60527	O60527 homo sapien
96	7	2.3	247	13	Q9IA22	Q9ia22 heterodontu	169	7	2.3	414	10	Q9M231	Q9m231 arabidopsis
97	7	2.3	249	5	Q9VPW8	Q9vpw8 drosophila	170	7	2.3	421	3	O13960	O13960 schizosacch
98	7	2.3	254	10	Q9LU18	Q9lu18 arabidopsis	171	7	2.3	422	2	P74896	P74896 thermus aqu
99	7	2.3	266	10	Q9SF29	Q9sf29 arabidopsis	172	7	2.3	423	1	Q9Y9U0	Q9y9u0 aeropyrum p
100	7	2.3	269	10	Q9S9T2	Q9s9t2 arabidopsis	173	7	2.3	425	1	O04007	O04007 dunaliella
101	7	2.3	271	5	O44077	O44077 chlamps nip	174	7	2.3	426	10	Q9LIY4	Q9liy4 oryza sativ
102	7	2.3	272	10	O65643	O65643 arabidopsis	175	7	2.3	429	2	O84249	O84249 chlamydia t
103	7	2.3	274	2	Q9ZDN1	Q9zdn1 rickettsia	176	7	2.3	432	2	Q9X804	Q9x804 streptomyc
104	7	2.3	274	5	Q9VFA0	Q9vfa0 drosophila	177	7	2.3	432	4	O75528	O75528 homo sapien
105	7	2.3	276	5	Q9VB04	Q9vba0 drosophila	178	7	2.3	433	2	Q9X8V6	Q9x8v6 streptomyc
106	7	2.3	276	10	Q9SBF4	Q9sbf4 arabidopsis	179	7	2.3	444	2	Q9ZF77	Q9zf77 bukkholderi
107	7	2.3	277	2	O24918	O24918 helicobacte	180	7	2.3	444	5	Q20584	Q20584 caenorhabdi
108	7	2.3	277	2	Q9ZMY0	Q9zmy0 helicobacte	181	7	2.3	445	10	Q9MBG7	Q9mbg7 arabidopsis
109	7	2.3	282	2	P72754	P72754 synecocyst	182	7	2.3	446	13	P79825	P79825 oncorhynch
110	7	2.3	283	4	Q9Z956	Q9z956 homo sapien	183	7	2.3	449	12	Q9PZ18	Q9pzt18 xestia c-ni
111	7	2.3	283	4	Q9UM65	Q9um65 homo sapien	184	7	2.3	453	2	O66700	O66700 aquifex aeo
112	7	2.3	283	6	Q9XS28	Q9xs28 cercopithe	185	7	2.3	457	5	Q19521	Q19521 caenorhabdi
113	7	2.3	285	10	Q9SU17	Q9sul7 arabidopsis	186	7	2.3	458	5	O60957	O60957 leishmania
114	7	2.3	286	10	O9LGS9	O9lgs9 oryza sativ	187	7	2.3	459	5	O45080	O45080 caenorhabdi
115	7	2.3	288	5	Q9VRL6	Q9vrl6 drosophila	188	7	2.3	467	4	Q9NPF5	Q9npf5 homo sapien
116	7	2.3	288	13	Q9YH08	Q9yh08 xenopus lae	189	7	2.3	468	11	Q9JII4	Q9ji44 mus musculu
117	7	2.3	290	2	O33434	O33434 paracoccus	190	7	2.3	469	5	Q9VFP3	Q9vfp3 drosophila
118	7	2.3	292	2	Q9X9X2	Q9x9x2 streptomyc	191	7	2.3	470	4	Q9UNP3	Q9unp3 homo sapien
119	7	2.3	293	5	O15988	O15988 patinopecte	192	7	2.3	472	5	Q20406	Q20406 caenorhabdi
120	7	2.3	294	5	Q9MAU5	Q9mau5 drosophila	193	7	2.3	479	4	Q9UPD0	Q9upd0 homo sapien
121	7	2.3	296	10	O80952	O80952 arabidopsis	194	7	2.3	480	3	Q9P874	Q9p874 candida alb
122	7	2.3	299	5	Q9NNF4	Q9nnf4 leishmania	195	7	2.3	483	11	Q9R255	Q9r255 mus musculu
123	7	2.3	303	2	O06545	O06545 mycobacteri	196	7	2.3	485	2	Q9LZL1	Q9lzl1 streptomyc
124	7	2.3	310	5	O16367	O16367 caenorhabdi	197	7	2.3	486	2	O86746	O86746 streptomyc
125	7	2.3	312	10	Q9SZW2	Q9szw2 arabidopsis	198	7	2.3	492	5	O18411	O18411 drosophila
126	7	2.3	314	5	O15987	O15987 patinopecte	199	7	2.3	495	4	Q9P2C2	Q9p2c2 homo sapien
127	7	2.3	315	4	O43597	O43597 homo sapien	200	7	2.3	498	1	Q9V2Y5	Q9v2y5 methanobac
128	7	2.3	315	11	Q9WUQ9	Q9wuq9 mus musculu	201	7	2.3	500	10	O24348	O24348 sorghum bic
129	7	2.3	315	11	Q9QXV8	Q9qxv8 mus musculu	202	7	2.3	501	2	Q9R2U8	Q9r2u8 streptomyc
130	7	2.3	321	10	O23662	O23662 arabidopsis	203	7	2.3	508	2	Q9KQR7	Q9kqr7 vibrio chol
131	7	2.3	324	5	Q9XY95	Q9xy95 lampetra fl	204	7	2.3	509	10	Q9SN03	Q9sn03 arabidopsis
132	7	2.3	327	6	O97491	O97491 ovis aries	205	7	2.3	512	4	Q9Y6P1	Q9y6p1 homo sapien
133	7	2.3	329	2	O45987	O45987 caulobacter	206	7	2.3	512	5	Q9V820	Q9v820 drosophila
134	7	2.3	330	5	Q9VWS4	Q9vws4 drosophila	207	7	2.3	513	2	O52237	O52237 myxococcus
135	7	2.3	332	2	O52695	O52695 rhodobacter	208	7	2.3	518	10	Q9SW67	Q9sw67 arabidopsis
136	7	2.3	332	10	Q9M8T8	Q9m8t8 arabidopsis	209	7	2.3	519	5	Q21928	Q21928 caenorhabdi
137	7	2.3	334	5	O9W3Y9	O9w3y9 drosophila	210	7	2.3	526	10	O22657	O22657 citrullus l
138	7	2.3	335	2	Q9P3X7	Q9p3x7 chlamydia m	211	7	2.3	531	5	Q9N8Y4	Q9n8y4 trypanosoma
139	7	2.3	335	2	Q9Z808	Q9z808 chlamydia p	212	7	2.3	532	5	Q9VVC8	Q9vvc8 drosophila
140	7	2.3	338	2	Q9Z954	Q9z954 chlamydia p	213	7	2.3	533	10	Q9M9K5	Q9m9k5 arabidopsis
141	7	2.3	338	5	Q9NAL9	Q9nal9 caenorhabdi	214	7	2.3	534	12	O36162	O36162 cyanophage
142	7	2.3	343	2	Q9K2C3	Q9k2c3 chlamydia p	215	7	2.3	537	5	Q9VP64	Q9vp64 drosophila
143	7	2.3	346	5	O44438	O44438 caenorhabdi	216	7	2.3	546	2	Q9ZC50	Q9zcs0 yersinia pe
144	7	2.3	349	3	O74947	O74947 schizosacch	217	7	2.3	548	4	Q9UP17	Q9up17 homo sapien
145	7	2.3	352	10	O22685	O22685 arabidopsis	218	7	2.3	554	5	O17459	O17459 schistosoma
146	7	2.3	356	4	Q9NUC0	Q9nuc0 homo sapien	219	7	2.3	557	3	O74537	O74537 schizosacch
147	7	2.3	360	11	Q9JL20	Q9jl20 mus musculu	220	7	2.3	557	10	Q9LQ10	Q9lq10 arabidopsis
148	7	2.3	361	4	Q9NZG2	Q9nzg2 homo sapien	221	7	2.3	560	4	O95443	O95443 homo sapien
149	7	2.3	362	11	Q9JL21	Q9jl21 mus musculu	222	7	2.3	560	5	Q9NNM7	Q9nmv7 leishmania
150	7	2.3	362	11	Q9JTP1	Q9jtp1 mus musculu	223	7	2.3	561	5	Q9V615	Q9v615 drosophila
151	7	2.3	364	3	O94124	O94124 rhizopus st	224	7	2.3	567	10	Q9SVL5	Q9svl5 arabidopsis
152	7	2.3	368	3	Q9P3B7	Q9p3b7 neurospora	225	7	2.3	568	5	Q9VY96	Q9vy96 drosophila
153	7	2.3	369	4	Q9PFS2	Q9pfs2 homo sapien	226	7	2.3	576	5	Q9VX72	Q9vx72 drosophila
154	7	2.3	369	10	O65249	O65249 arabidopsis	227	7	2.3	583	5	Q9VNZ5	Q9vnz5 drosophila
155	7	2.3	374	2	Q9RJF6	Q9rjf6 streptomyc	228	7	2.3	593	12	Q9IBL3	Q9ibl3 budgerigar
156	7	2.3	381	8	P92341	P92341 restrepla m	229	7	2.3	600	2	P74569	P74569 synecocyst
157	7	2.3	381	10	Q9LD39	Q9ld39 arabidopsis	230	7	2.3	601	10	O23522	O23522 arabidopsis
158	7	2.3	382	5	P91703	P91703 drosophila	231	7	2.3	603	10	O22787	O22787 arabidopsis
159	7	2.3	385	2	Q9X4V1	Q9x4v1 streptococ	232	7	2.3	607	10	Q9LGS5	Q9lgs5 oryza sativ
160	7	2.3	391	3	Q12132	Q12132 saccharomyc	233	7	2.3	623	2	O30589	O30589 streptomyc
161	7	2.3	392	4	Q9URF6	Q9urf6 homo sapien	234	7	2.3	623	5	Q9VKK1	Q9vkk1 drosophila
162	7	2.3	393	1	Q9YBB6	Q9ybb6 aeropyrum p	235	7	2.3	630	3	O13870	O13870 schizosacch
163	7	2.3	401	12	O11375	O11375 molluscum c	236	7	2.3	640	10	O23650	O23650 arabidopsis
164	7	2.3	406	12	O98312	O98312 molluscum c	237	7	2.3	641	5	O01667	O01667 drosophila
165	7	2.3	407	2	Q9ZCV6	Q9zcv6 rickettsia	238	7	2.3	641	5	Q9VU01	Q9vu01 drosophila

239	7	2.3	642	10	Q42590	Q42590 arabidopsis	312	7	2.3	1155	10	Q9SQ12	Q9sq12 arabidopsis
240	7	2.3	642	10	Q9L266	Q9l266 arabidopsis	313	7	2.3	1157	5	Q9VP22	Q9vp22 drosophila
241	7	2.3	657	5	Q9VI63	Q9vi63 drosophila	314	7	2.3	1167	10	O64381	O64381 arabidopsis
242	7	2.3	658	10	Q9SGK3	Q9sgk3 arabidopsis	315	7	2.3	1173	10	Q9S7N0	Q9s7n0 arabidopsis
243	7	2.3	659	3	Q12256	Q12256 saccharomyc	316	7	2.3	1182	5	Q9VH60	Q9vh60 drosophila
244	7	2.3	662	5	Q9VFD5	Q9vfd5 drosophila	317	7	2.3	1198	5	O17720	O17720 caenorhabdi
245	7	2.3	665	5	Q23823	Q23823 calliphora	318	7	2.3	1217	5	Q17240	Q17240 bombyx mori
246	7	2.3	671	5	P81538	P81538 drosophila	319	7	2.3	1225	4	Q9UJ98	Q9uj98 homo sapien
247	7	2.3	688	10	Q9SVd7	Q9svd7 arabidopsis	320	7	2.3	1256	5	Q9VG16	Q9vg16 drosophila
248	7	2.3	694	5	Q9VS05	Q9vs05 drosophila	321	7	2.3	1268	2	Q9KZL1	Q9kzl1 streptomyce
249	7	2.3	695	5	Q9VU44	Q9vu44 drosophila	322	7	2.3	1279	5	O97005	O97005 leishmania
250	7	2.3	697	10	P93203	P93203 lycopersico	323	7	2.3	1289	11	Q9WTR2	Q9wtr2 mus musculu
251	7	2.3	700	5	Q9W3S2	Q9w3s2 drosophila	324	7	2.3	1297	4	O15025	O15025 homo sapien
252	7	2.3	704	8	Q32143	Q32143 danthionlops	325	7	2.3	1318	12	P90493	P90493 herpes simp
253	7	2.3	705	5	Q9NBW1	Q9nbw1 drosophila	326	7	2.3	1320	5	Q9W0Q1	Q9w0q1 drosophila
254	7	2.3	708	2	Q46444	Q46444 comamanas t	327	7	2.3	1323	5	Q9NHX4	Q9nhx4 drosophila
255	7	2.3	726	11	Q9J125	Q9j125 mus musculu	328	7	2.3	1339	5	Q26048	Q26048 pacifastacu
256	7	2.3	729	13	Q9O971	Q9o971 gallus gall	329	7	2.3	1400	5	O9NKS7	O9nks7 leishmania
257	7	2.3	733	13	O73897	O73897 gallus gall	330	7	2.3	1440	2	O69498	O69498 mycobacteri
258	7	2.3	734	10	O48620	O48620 populus tri	331	7	2.3	1469	4	O9Y6P7	O9y6p7 homo sapien
259	7	2.3	735	10	O48619	O48619 populus tri	332	7	2.3	1487	5	O15843	O15843 leishmania
260	7	2.3	737	5	Q23985	Q23985 drosophila	333	7	2.3	1521	5	Q9VZS2	Q9vzs2 drosophila
261	7	2.3	738	5	Q9W3Z1	Q9w3z1 drosophila	334	7	2.3	1543	5	Q9W406	Q9w406 drosophila
262	7	2.3	741	4	O9NSZ1	O9nsz1 homo sapien	335	7	2.3	1548	10	O65531	O65531 arabidopsis
263	7	2.3	743	5	O00903	O00903 leishmania	336	7	2.3	1560	5	Q9VRY4	Q9vry4 drosophila
264	7	2.3	754	4	O15310	O15310 homo sapien	337	7	2.3	1611	2	O33957	O33957 streptomyce
265	7	2.3	757	4	Q9P237	Q9p237 homo sapien	338	7	2.3	1681	11	O62467	O62467 mus musculu
266	7	2.3	793	5	Q9VWZ7	Q9vwz7 drosophila	339	7	2.3	1809	5	Q9N4C0	Q9n4c0 caenorhabdi
267	7	2.3	795	5	O62006	O62006 brachiosteo	340	7	2.3	1844	5	Q9V9S7	Q9v9s7 drosophila
268	7	2.3	798	5	O9U4H1	O9u4h1 drosophila	341	7	2.3	1846	5	O9VWN8	Q9vwn8 drosophila
269	7	2.3	798	11	O88411	O88411 mus musculu	342	7	2.3	1907	5	Q9U5D8	Q9u5d8 plautia sta
270	7	2.3	798	11	O54795	O54795 mus musculu	343	7	2.3	1937	5	Q9W003	Q9w003 drosophila
271	7	2.3	802	5	Q9VL43	Q9vl43 drosophila	344	7	2.3	2006	5	Q9VWF2	Q9vzf2 drosophila
272	7	2.3	803	4	O15022	O15022 homo sapien	345	7	2.3	2110	5	Q9VRA6	Q9vra6 drosophila
273	7	2.3	810	5	O45031	O45031 drosophila	346	7	2.3	2185	3	Q12721	Q12721 ustilago ma
274	7	2.3	813	5	P90538	P90538 dictyostei	347	7	2.3	2187	11	P70670	P70670 mus musculu
275	7	2.3	832	2	Q9LA49	Q9la49 treponema p	348	7	2.3	2232	5	P91365	P91365 caenorhabdi
276	7	2.3	832	2	Q9LA48	Q9la48 treponema s	349	7	2.3	2304	5	Q9W548	Q9w548 drosophila
277	7	2.3	837	2	O68454	O68454 treponema p	350	7	2.3	2422	5	O46083	O46083 drosophila
278	7	2.3	837	2	Q9LA51	Q9la51 treponema p	351	7	2.3	2539	12	O9J7C1	O9j7cl porcine rep
279	7	2.3	837	2	Q9LA50	Q9la50 treponema p	352	7	2.3	2571	2	O87704	O87704 bacillus su
280	7	2.3	837	12	Q9JRL6	Q9jrl6 treponema p	353	7	2.3	2921	5	O9N973	O9n973 leishmania
281	7	2.3	843	5	Q9V866	Q9vzt9 hepatitis b	354	7	2.3	3080	5	O9VRY3	O9vry3 drosophila
282	7	2.3	843	5	Q9V866	Q9v866 drosophila	355	7	2.3	3582	2	O66069	O66069 bacillus li
283	7	2.3	843	12	O81107	O81107 hepatitis b	356	7	2.3	3583	2	O45295	O45295 bacillus li
284	7	2.3	849	3	P87107	P87107 saccharomyc	357	7	2.3	4005	4	Q13744	Q13744 homo sapien
285	7	2.3	853	2	O83346	O83346 treponema p	358	7	2.3	4599	4	O9NZR2	O9nizr2 homo sapien
286	7	2.3	856	5	Q9V706	Q9v706 drosophila	359	7	2.3	4599	11	Q9J118	Q9j118 mus musculu
287	7	2.3	865	5	O18395	O18395 drosophila	360	6	2.0	20	12	Q9PXE4	Q9pxe4 foot-and-mo
288	7	2.3	866	5	O9NBW9	O9nbw9 drosophila	361	6	2.0	23	12	O87081	O87081 pseudorabie
289	7	2.3	870	6	Q9XTA4	Q9xta4 bos taurus	362	6	2.0	25	11	O9JXY4	O9jxy4 mus musculu
290	7	2.3	873	5	Q9Y076	Q9y076 leishmania	363	6	2.0	36	6	Q28316	Q28316 capra hircu
291	7	2.3	914	2	O33964	O33964 streptomyce	364	6	2.0	39	2	O9PFD3	Q9pfd3 xylella fas
292	7	2.3	928	5	O18696	O18696 caenorhabdi	365	6	2.0	40	2	O46498	O46498 desulfovibr
293	7	2.3	944	4	Q14163	Q14163 homo sapien	366	6	2.0	43	4	O14601	O14601 homo sapien
294	7	2.3	954	5	Q9VFE5	Q9vfe5 drosophila	367	6	2.0	44	12	O9YR19	O9yrl9 infectious
295	7	2.3	968	5	O9N990	O9n990 leishmania	368	6	2.0	50	2	O9XBB5	O9xbbs staphylococ
296	7	2.3	971	4	O75432	O75432 homo sapien	369	6	2.0	52	5	O9VSL1	Q9vsl1 drosophila
297	7	2.3	971	4	O9UP99	O9up99 homo sapien	370	6	2.0	56	1	O9YFF5	Q9yff5 aeropyrum p
298	7	2.3	971	13	Q9PTU3	Q9ptu3 pagrus majo	371	6	2.0	56	10	O9LWK6	O9lwk6 oryza sativ
299	7	2.3	972	10	Q9ZPY7	Q9zpy7 arabidopsis	372	6	2.0	56	11	O61725	O61725 mus musculu
300	7	2.3	1034	5	O16015	O16015 drosophila	373	6	2.0	61	2	Q9X4G6	Q9x4g6 vibrio chol
301	7	2.3	1037	5	O9VZ53	Q9vz53 drosophila	374	6	2.0	62	6	O29359	O29359 sus scrofa
302	7	2.3	1054	5	Q9N9M5	Q9n9m5 leishmania	375	6	2.0	62	12	O89183	O89183 variola vir
303	7	2.3	1057	2	Q9K294	Q9k294 drosophila	376	6	2.0	63	5	O9VBH1	Q9vbm1 drosophila
304	7	2.3	1063	2	O9KDU5	Q9kdu5 bacillus ha	377	6	2.0	73	2	O50014	O50014 mycobacteri
305	7	2.3	1070	10	O80545	O80545 arabidopsis	378	6	2.0	76	4	O00662	O00662 homo sapien
306	7	2.3	1070	10	O9SV36	O9sv36 arabidopsis	379	6	2.0	76	6	P79186	P79186 macaca fusc
307	7	2.3	1126	13	O9PTN6	O9ptn6 cyprinus ca	380	6	2.0	76	6	P79796	P79796 hylobates c
308	7	2.3	1134	4	O9PIF7	O9pif7 homo sapien	381	6	2.0	76	10	O40497	O40497 nicotiana t
309	7	2.3	1134	4	O9NPG3	O9npg3 homo sapien	382	6	2.0	76	10	O9M690	O9m690 gossypioide
310	7	2.3	1135	13	O9YHD3	O9yhd3 xenopus lae	383	6	2.0	77	2	O9L459	Q9l459 propionibac
311	7	2.3	1150	3	Q99129	Q99129 ustilago ma	384	6	2.0	80	2	Q9ZAX9	Q9zax9 sphingomona

385	6	2.0	80	2	Q9S240	Q9S240 streptomyce	458	6	2.0	118	10	Q9LK37	Q9LK37 arabidopsis
386	6	2.0	80	5	Q23341	Q23341 caenorhabdi	459	6	2.0	119	5	Q9V401	Q9V401 drosophila
387	6	2.0	80	5	P81793	P81793 phoneutria	460	6	2.0	119	8	Q9MFA6	Q9MFA6 beta vulgar
388	6	2.0	80	5	Q9V542	Q9V542 drosophila	461	6	2.0	120	5	Q9VNS7	Q9VNS7 drosophila
389	6	2.0	81	10	Q9M645	Q9M645 oryza sativ	462	6	2.0	120	10	Q39938	Q39938 helianthus
390	6	2.0	81	12	Q88513	Q88513 porcine tra	463	6	2.0	120	12	Q98563	Q98563 paramecium
391	6	2.0	84	10	Q9SHW8	Q9SHW8 arabidopsis	464	6	2.0	122	2	Q9RJX8	Q9RJX8 streptomyce
392	6	2.0	84	5	Q18084	Q18084 caenorhabdi	465	6	2.0	122	10	Q9ZRUI	Q9ZRUI dactylis gl
393	6	2.0	84	5	P91899	P91899 bombyx mori	466	6	2.0	122	10	Q9ZRUI	Q9ZRUI cydonod dac
394	6	2.0	86	2	Q50112	Q50112 mycobacteri	467	6	2.0	122	10	Q9LNQ3	Q9LNQ3 arabidopsis
395	6	2.0	86	4	Q9P2X8	Q9P2X8 homo sapien	468	6	2.0	123	5	Q21019	Q21019 caenorhabdi
396	6	2.0	86	12	Q9QPR8	Q9QPR8 hepatitis c	469	6	2.0	123	6	Q9MZI7	Q9MZI7 oryctolagus
397	6	2.0	94	5	Q23798	Q23798 chironomus	470	6	2.0	123	10	Q42253	Q42253 arabidopsis
398	6	2.0	94	11	Q08937	Q08937 mus spretus	471	6	2.0	124	1	Q9YC68	Q9YC68 aeropyrum p
399	6	2.0	95	1	Q58126	Q58126 methanococ	472	6	2.0	124	5	Q94045	Q94045 caenorhabdi
400	6	2.0	95	10	Q42260	Q42260 arabidopsis	473	6	2.0	124	12	Q66898	Q66898 four corner
401	6	2.0	96	4	Q9Y3X1	Q9Y3X1 homo sapien	474	6	2.0	125	5	Q05757	Q05757 schistocerc
402	6	2.0	96	5	Q9VN78	Q9VN78 drosophila	475	6	2.0	125	5	Q21539	Q21539 caenorhabdi
403	6	2.0	96	5	Q9NM43	Q9NM43 leishmania	476	6	2.0	126	2	P73010	P73010 synechocyst
404	6	2.0	98	8	Q21541	Q21541 bolomys lac	477	6	2.0	126	5	Q27011	Q27011 tenebrio mo
405	6	2.0	98	8	Q9MEI2	Q9MEI2 lama guanac	478	6	2.0	126	10	Q9LPC9	Q9LPC9 pseudomonas
406	6	2.0	98	10	Q65086	Q65086 picea maria	479	6	2.0	127	2	Q9PK93	Q9PK93 chlamydia m
407	6	2.0	99	2	Q47546	Q47546 escherichia	480	6	2.0	127	2	Q9PK93	Q9PK93 myxococcus
408	6	2.0	99	8	Q9MGD1	Q9MGD1 penaeus mon	481	6	2.0	128	2	Q50896	Q50896 thermotoga
409	6	2.0	100	8	Q47541	Q47541 chlamydomon	482	6	2.0	128	2	Q9WZK5	Q9WZK5 bos taurus
410	6	2.0	101	2	Q9R7Q9	Q9R7Q9 escherichia	483	6	2.0	128	6	Q28195	Q28195 arabidopsis
411	6	2.0	101	2	Q9LBN1	Q9LBN1 uncultured	484	6	2.0	128	10	Q9T068	Q9T068 arabidopsis
412	6	2.0	102	2	Q86496	Q86496 streptomyce	485	6	2.0	129	1	Q9YBX0	Q9YBX0 aeropyrum p
413	6	2.0	102	3	Q08736	Q08736 saccharomyc	486	6	2.0	129	4	Q9XA01	Q9XA01 streptomyce
414	6	2.0	102	8	Q9MF81	Q9MF81 beta vulgar	487	6	2.0	129	4	Q9NT15	Q9NT15 homo sapien
415	6	2.0	102	10	Q9SG43	Q9SG43 arabidopsis	488	6	2.0	129	5	Q44535	Q44535 drosophila
416	6	2.0	102	10	Q9S738	Q9S738 lycopersico	489	6	2.0	129	5	Q9VNS6	Q9VNS6 drosophila
417	6	2.0	102	12	Q68358	Q68358 hepatitis c	490	6	2.0	129	6	Q46541	Q46541 ovis aries
418	6	2.0	103	2	Q50015	Q50015 mycobacteri	491	6	2.0	129	10	Q49715	Q49715 arabidopsis
419	6	2.0	103	4	Q9NU29	Q9NU29 homo sapien	492	6	2.0	130	1	Q9UWU6	Q9UWU6 sulfolobus
420	6	2.0	104	2	Q9LC05	Q9LC05 staphylococ	493	6	2.0	130	4	Q9S540	Q9S540 homo sapien
421	6	2.0	104	10	Q9LMA2	Q9LMA2 arabidopsis	494	6	2.0	130	4	Q9NZ82	Q9NZ82 homo sapien
422	6	2.0	105	10	Q42254	Q42254 arabidopsis	495	6	2.0	130	10	Q9SECC	Q9SECC pimptinella
423	6	2.0	105	10	Q9SAB8	Q9SAB8 arabidopsis	496	6	2.0	130	12	Q9QU32	Q9QU32 ttv-like mi
424	6	2.0	105	12	Q56310	Q56310 egyptian su	497	6	2.0	130	12	Q9QU29	Q9QU29 ttv-like mi
425	6	2.0	105	12	Q9YXE2	Q9YXE2 egyptian su	498	6	2.0	131	2	Q9PNK8	Q9PNK8 campylobact
426	6	2.0	105	12	Q9YXD8	Q9YXD8 egyptian su	499	6	2.0	131	10	Q9M595	Q9M595 elaeis guin
427	6	2.0	105	12	Q9YXD5	Q9YXD5 egyptian su	500	6	2.0	131	10	Q9LSX0	Q9LSX0 arabidopsis
428	6	2.0	105	12	Q9YH25	Q9YH25 egyptian su	501	6	2.0	132	2	Q06662	Q06662 shigella fi
429	6	2.0	106	1	Q9YBT1	Q9YBT1 aeropyrum p	502	6	2.0	132	4	Q9Y3Y7	Q9Y3Y7 homo sapien
430	6	2.0	106	10	Q41299	Q41299 sorghum bic	503	6	2.0	133	1	Q9YGO3	Q9YGO3 aeropyrum p
431	6	2.0	107	2	Q06351	Q06351 mycobacteri	504	6	2.0	133	2	Q67824	Q67824 aquifex aeo
432	6	2.0	108	1	Q98919	Q98919 methanococ	505	6	2.0	134	13	P70072	P70072 gallus gall
433	6	2.0	108	4	Q9NTL7	Q9NTL7 homo sapien	506	6	2.0	135	5	Q9VYT9	Q9VYT9 drosophila
434	6	2.0	108	5	Q9NM03	Q9NM03 leishmania	507	6	2.0	135	8	Q47945	Q47945 lumbricus r
435	6	2.0	109	2	Q68643	Q68643 pseudomonas	508	6	2.0	135	10	Q39792	Q39792 gossypium h
436	6	2.0	109	10	Q65164	Q65164 mesembryant	509	6	2.0	135	10	Q9LGY6	Q9LGY6 oryza sativ
437	6	2.0	110	2	Q9KYV8	Q9KYV8 streptomyce	510	6	2.0	135	12	Q68199	Q68199 hepatitis c
438	6	2.0	110	5	Q9W5Y8	Q9W5Y8 drosophila	511	6	2.0	136	2	Q9RSG1	Q9RSG1 deinococcus
439	6	2.0	110	11	Q9RIQ1	Q9RIQ1 mus musculu	512	6	2.0	137	10	Q40043	Q40043 hordeum vul
440	6	2.0	111	2	Q87762	Q87762 lactococcus	513	6	2.0	137	10	Q9S839	Q9S839 arabidopsis
441	6	2.0	111	2	Q9LBM8	Q9LBM8 uncultured	514	6	2.0	139	2	P97110	P97110 escherichia
442	6	2.0	112	2	Q9LBP9	Q9LBP9 chemosynthe	515	6	2.0	139	2	Q54289	Q54289 salmonella
443	6	2.0	112	2	Q9LBN4	Q9LBN4 uncultured	516	6	2.0	139	2	P95002	P95002 mycobacteri
444	6	2.0	112	2	Q9LBM6	Q9LBM6 uncultured	517	6	2.0	139	10	P93179	P93179 hordeum vul
445	6	2.0	112	2	Q9LBM3	Q9LBM3 uncultured	518	6	2.0	139	10	Q9LTC5	Q9LTC5 arabidopsis
446	6	2.0	112	2	Q9LBM2	Q9LBM2 uncultured	519	6	2.0	140	2	Q83445	Q83445 treponema p
447	6	2.0	112	5	Q16119	Q16119 tenebrio mo	520	6	2.0	140	10	Q9LPP3	Q9LPP3 arabidopsis
448	6	2.0	112	10	Q9M296	Q9M296 arabidopsis	521	6	2.0	140	11	Q64145	Q64145 rattus norv
449	6	2.0	113	2	Q9LBO1	Q9LBO1 uncultured	522	6	2.0	141	2	Q9S414	Q9S414 actinobacil
450	6	2.0	113	2	Q9LBO0	Q9LBO0 uncultured	523	6	2.0	141	4	Q9NXX4	Q9NXX4 homo sapien
451	6	2.0	113	2	Q9LBP8	Q9LBP8 uncultured	524	6	2.0	141	5	Q9VZV5	Q9VZV5 drosophila
452	6	2.0	113	4	Q12925	Q12925 homo sapien	525	6	2.0	141	10	Q9SP80	Q9SP80 hordeum vul
453	6	2.0	113	10	Q9LDB4	Q9LDB4 arabidopsis	526	6	2.0	141	10	Q9LWK3	Q9LWK3 oryza sativ
454	6	2.0	117	2	P71212	P71212 escherichia	527	6	2.0	142	2	Q9RY90	Q9RY90 deinococcus
455	6	2.0	117	5	Q9VNS4	Q9VNS4 drosophila	528	6	2.0	143	10	Q80782	Q80782 arabidopsis
456	6	2.0	118	5	Q9ULK5	Q9ULK5 drosophila	529	6	2.0	143	10	Q9SBI9	Q9SBI9 hordeum vul
457	6	2.0	118	6	Q9TVA2	Q9TVA2 sus scrofa	530	6	2.0	144	2	Q84950	Q84950 salmonella

531	6	2.0	144	2	Q9JXP9	Q9jpx9 bacillus ha	604	169	10	Q9SE91	Q9se91 brassica ol
532	6	2.0	144	11	Q9JJD0	Q9jid0 mus musculus	605	169	10	Q9LWD1	Q9lwd1 oryza sativ
533	6	2.0	145	2	Q97A04	Q97a04 mycobacteri	606	170	2	Q9RTL6	Q9rtl6 deinococcus
534	6	2.0	145	2	Q9RYG6	Q9ryg6 deinococcus	607	170	3	Q9P4B4	Q9p4b4 letharia vu
535	6	2.0	145	2	Q9RIK1	Q9rik1 streptococc	608	170	10	Q9LX15	Q9lxi15 arabidopsis
536	6	2.0	145	3	Q90233	Q90233 arthrobotry	609	170	10	Q9LXV6	Q9lxv6 arabidopsis
537	6	2.0	145	10	Q93791	Q93791 gossypium h	610	170	10	Q9LVV6	Q9lvv6 arabidopsis
538	6	2.0	146	4	Q9UNX1	Q9unx1 homo sapien	611	171	2	Q9JVS5	Q9jvs5 neisseria m
539	6	2.0	146	10	Q92TR3	Q92tr3 hordeum vul	612	171	5	Q9V9S9	Q9v9s9 drosophila
540	6	2.0	147	5	Q9VCD9	Q9vcd9 drosophila	613	171	11	Q9JLF0	Q9jlf0 rattus norv
541	6	2.0	148	2	Q93397	Q93397 mycobacteri	614	172	2	Q9X879	Q9x879 streptomyce
542	6	2.0	148	4	Q9NVH7	Q9nvh7 homo sapien	615	172	2	Q9LLK0	Q9llk0 streptomyce
543	6	2.0	148	5	Q9W4G0	Q9w4g0 drosophila	616	172	4	Q9NPA9	Q9npa9 homo sapien
544	6	2.0	149	1	Q9Y9V4	Q9y9v4 aeropyrum p	617	172	5	Q96933	Q96933 tetrahymena
545	6	2.0	149	2	Q9RZA5	Q9rza5 deinococcus	618	173	4	Q9UP48	Q9up48 homo sapien
546	6	2.0	149	4	Q9NXJ8	Q9nxj8 homo sapien	619	173	10	Q9LYB4	Q9lyb4 arabidopsis
547	6	2.0	149	5	Q92I091	Q92i091 caenorhabdi	620	174	1	Q30090	Q30090 archaeoglob
548	6	2.0	149	5	Q9NB84	Q9nb84 agrotis ips	621	174	2	Q96001	Q96001 synechococc
549	6	2.0	149	11	Q9WUX6	Q9wux6 uranomyces ru	622	174	2	Q9RI63	Q9ri63 streptomyce
550	6	2.0	150	5	Q18551	Q18551 lytechinus	623	174	5	Q90282	Q90282 caenorhabdi
551	6	2.0	150	10	Q41601	Q41601 triticum tu	624	175	1	Q29754	Q29754 archaeoglob
552	6	2.0	150	10	Q81464	Q81464 arabidopsis	625	175	1	Q57788	Q57788 pyrococcus
553	6	2.0	151	2	Q93879	Q93879 streptomyce	626	175	2	Q73596	Q73596 synechocyst
554	6	2.0	151	5	Q18728	Q18728 caenorhabdi	627	175	2	Q53808	Q53808 mycobacteri
555	6	2.0	151	10	Q40186	Q40186 lemna gibba	628	175	5	Q9VTA6	Q9vta6 drosophila
556	6	2.0	152	5	Q06777	Q06777 haematobia	629	175	11	Q54926	Q54926 mus musculu
557	6	2.0	152	12	Q9YUH8	Q9yuh8 rhesus cyto	630	175	11	Q9RIQ0	Q9riq0 mus musculu
558	6	2.0	153	2	Q9X8K2	Q9x8k2 streptomyce	631	176	5	Q9V734	Q9v734 drosophila
559	6	2.0	154	2	Q41287	Q41287 sorghum bic	632	176	10	Q9SV01	Q9svu1 arabidopsis
560	6	2.0	155	2	Q9WTF8	Q9wtf8 pseudomonas	633	176	10	Q9LIJ0	Q9lij0 arabidopsis
561	6	2.0	155	2	Q9S2N9	Q9s2n9 streptomyce	634	177	2	Q9LAZ4	Q9laz4 streptomyce
562	6	2.0	155	2	Q9L0P6	Q9l0p6 streptomyce	635	177	10	Q9M9H3	Q9m9h3 arabidopsis
563	6	2.0	155	3	Q90569	Q90569 saccharomyc	636	177	12	Q66264	Q66264 colorado ti
564	6	2.0	155	5	Q9VCF9	Q9vcf9 drosophila	637	177	12	Q41996	Q41996 maize rayad
565	6	2.0	155	10	Q9ZTR7	Q9ztr7 hordeum vul	638	178	5	Q22848	Q22848 caenorhabdi
566	6	2.0	155	10	Q9SPA9	Q9spa9 hordeum vul	639	178	5	Q9XWH9	Q9xwh9 caenorhabdi
567	6	2.0	155	12	Q11842	Q11842 abutilon mo	640	179	9	Q37954	Q37954 lactococcus
568	6	2.0	156	4	Q60888	Q60888 homo sapien	641	179	11	Q9JKW9	Q9jkw9 mus musculu
569	6	2.0	156	10	Q64940	Q64940 lophopyrum	642	180	6	Q62822	Q62822 bubalus bub
570	6	2.0	157	5	Q9W482	Q9w482 drosophila	643	180	10	Q82297	Q82297 arabidopsis
571	6	2.0	157	10	Q48621	Q48621 solanum com	644	180	10	Q9SSY1	Q9ssy1 cucumis sat
572	6	2.0	158	5	Q9NA77	Q9na77 caenorhabdi	645	180	12	Q66263	Q66263 colorado ti
573	6	2.0	158	10	Q41600	Q41600 triticum tu	646	180	13	Q90045	Q90045 xenopus lae
574	6	2.0	159	2	Q9S2B8	Q9s2b8 streptomyce	647	181	2	Q82986	Q82986 bacillus sp
575	6	2.0	159	2	Q9RXP4	Q9rxp4 deinococcus	648	181	10	Q24533	Q24533 vicia faba
576	6	2.0	159	2	Q9KXV2	Q9kxv2 streptomyce	649	181	10	Q9ZTR4	Q9ztr4 hordeum vul
577	6	2.0	160	5	Q9N7S5	Q9n7s5 leishmania	650	182	10	Q9SGP4	Q9sgp4 arabidopsis
578	6	2.0	161	10	Q93706	Q93706 dahlia sp.	651	182	13	Q9SXP4	Q9sxp4 gallus gall
579	6	2.0	161	10	Q9M242	Q9m242 arabidopsis	652	183	4	Q9NX31	Q9nx31 homo sapien
580	6	2.0	161	12	Q94531	Q94531 pseudorabie	653	185	2	Q51143	Q51143 nitrosolobu
581	6	2.0	161	12	Q9QNC2	Q9qnc2 cucumber gr	654	185	5	Q9NNG6	Q9nnng leishmania
582	6	2.0	162	5	Q9W309	Q9w309 drosophila	655	185	10	Q96261	Q96261 arabidopsis
583	6	2.0	162	11	P70598	P70598 rattus norv	656	185	10	Q9SGP3	Q9sgp3 arabidopsis
584	6	2.0	163	2	Q53744	Q53744 mycobacteri	657	185	12	Q66262	Q66262 colorado ti
585	6	2.0	163	5	Q18679	Q18679 trypanosoma	658	186	1	Q9UZJ3	Q9uzj3 pyrococcus
586	6	2.0	163	10	Q9LTK7	Q9ltk7 arabidopsis	659	186	1	Q9UYD0	Q9uyd0 chlamydia m
587	6	2.0	164	2	Q51260	Q51260 borrelia bu	660	186	2	Q9PL86	Q9pl86 chlamydia m
588	6	2.0	165	1	Q26904	Q26904 methanobact	661	186	2	Q9KQ82	Q9kq82 vibrio chol
589	6	2.0	165	10	Q40651	Q40651 oryza sativ	662	186	5	Q44137	Q44137 caenorhabdi
590	6	2.0	166	2	Q9PKS0	Q9pks0 chlamydia m	663	186	5	Q45491	Q45491 caenorhabdi
591	6	2.0	166	3	Q13522	Q13522 saccharomyc	664	187	2	P93414	P93414 oryza sativ
592	6	2.0	166	5	Q9NLV3	Q9nlv3 leishmania	665	187	2	Q50944	Q50944 borrelia bu
593	6	2.0	166	10	Q23957	Q23957 glycine max	666	187	12	Q11329	Q11329 molluscum c
594	6	2.0	166	10	Q9SR81	Q9sr81 arabidopsis	667	189	2	Q98866	Q98866 streptomyce
595	6	2.0	166	12	Q55779	Q55779 hendra viru	668	189	5	Q17467	Q17467 caenorhabdi
596	6	2.0	167	2	Q9RKA3	Q9rka3 streptomyce	669	189	5	Q9TVU3	Q9tvu3 caenorhabdi
597	6	2.0	167	5	Q9WV15	Q9wv15 drosophila	670	190	2	Q9ZBE9	Q9zbe9 streptomyce
598	6	2.0	167	8	Q04342	Q04342 helianthus	671	190	2	Q9JZ44	Q9jza4 neisseria m
599	6	2.0	168	2	Q06085	Q06085 mycobacteri	672	190	2	Q9JUC7	Q9juc7 neisseria m
600	6	2.0	168	5	Q9NNH5	Q9nnh5 leishmania	673	190	10	Q9M3V9	Q9m3v9 oryza sativ
601	6	2.0	168	12	Q84423	Q84423 paramecium	674	191	2	Q9R7Q8	Q9r7q8 escherichia
602	6	2.0	169	2	Q07429	Q07429 mycobacteri	675	191	6	Q99491	Q99491 homo sapien
603	6	2.0	169	2	Q9K0P7	Q9k0p7 neisseria m	676	191	6	Q9TS03	Q9ts03 ovine arles
										Q38966	Q38966 arabidopsis

677	6	2.0	191	10	Q9SPA5	Q9spa5 hordeum vul	750	6	2.0	205	10	Q9SCS8	Q9scs8 arabidopsis
678	6	2.0	192	5	Q9VZG2	Q9vzg2 drosophila	751	6	2.0	206	2	Q45929	Q45929 coxiella bu
679	6	2.0	192	10	Q65416	Q65416 arabidopsis	752	6	2.0	206	2	Q45854	Q45854 coxiella bu
680	6	2.0	193	2	Q33334	Q33334 mycobacteri	753	6	2.0	206	2	Q9KRH5	Q9krh5 vibrio chol
681	6	2.0	193	6	Q28530	Q28530 macropus ro	754	6	2.0	206	4	Q9P0R8	Q9p0r8 homo sapien
682	6	2.0	194	1	Q9YB78	Q9yb78 aeropyrum p	755	6	2.0	206	4	Q9N284	Q9n284 homo sapien
683	6	2.0	194	2	P74432	P74432 synechocyst	756	6	2.0	206	5	Q22415	Q22415 caenorhabdi
684	6	2.0	194	3	Q06854	Q06854 saccharomyc	757	6	2.0	206	10	Q9LDB9	Q9ldh9 arabidopsis
685	6	2.0	195	2	Q9RJ36	Q9rj36 streptomyc	758	6	2.0	206	13	Q42606	Q42606 brachydanio
686	6	2.0	195	2	Q9RFJ2	Q9rfj2 streptococc	759	6	2.0	208	10	Q9SS31	Q9ss31 arabidopsis
687	6	2.0	195	2	Q9KNL2	Q9knl2 vibrio chol	760	6	2.0	209	10	Q9LMP8	Q9lmp8 oryza sativ
688	6	2.0	195	4	Q9ULZ3	Q9ulz3 homo sapien	761	6	2.0	210	2	Q9RNU2	Q9rnu2 streptococc
689	6	2.0	195	5	Q9VBE8	Q9vbe8 drosophila	762	6	2.0	210	3	Q9P559	Q9p559 neurospora
690	6	2.0	196	2	Q84231	Q84231 chlamydia t	763	6	2.0	211	2	Q9X6X8	Q9x6x8 streptococc
691	6	2.0	196	13	Q9YH31	Q9yh31 notophthalm	764	6	2.0	211	5	Q61575	Q61575 pristiionchu
692	6	2.0	197	1	Q28693	Q28693 archaeeoglob	765	6	2.0	211	5	Q9NLI9	Q9nli9 leishmania
693	6	2.0	197	2	Q06983	Q06983 bacillus su	766	6	2.0	212	1	Q28986	Q28986 archaeeoglob
694	6	2.0	197	2	Q9X8T3	Q9x8t3 streptomyc	767	6	2.0	212	2	Q9P898	Q9p898 xylella fas
695	6	2.0	197	2	Q9X728	Q9x728 escherichia	768	6	2.0	213	2	Q9RXD6	Q9rxd6 deinococcus
696	6	2.0	197	10	Q9LMB4	Q9lmb4 arabidopsis	769	6	2.0	213	2	Q9RFI6	Q9rfi6 streptococc
697	6	2.0	198	2	Q47300	Q47300 escherichia	770	6	2.0	213	6	Q9TUC8	Q9tuc8 monodelphis
698	6	2.0	198	2	Q9RJ21	Q9rj21 chlamydia m	771	6	2.0	213	10	Q9SCS7	Q9scs7 arabidopsis
699	6	2.0	198	10	Q9LRN9	Q9lrn9 arabidopsis	772	6	2.0	214	3	Q94190	Q94190 emericella
700	6	2.0	199	4	Q9Y4L8	Q9y4l8 homo sapien	773	6	2.0	216	2	Q56671	Q56671 vibrio chol
701	6	2.0	199	4	Q9NWT8	Q9nwt8 homo sapien	774	6	2.0	216	5	Q10910	Q10910 caenorhabdi
702	6	2.0	199	6	Q9T548	Q9ts48 ovis aries	775	6	2.0	216	5	Q9VFL3	Q9vfl3 drosophila
703	6	2.0	199	10	Q93386	Q93386 brassica ol	776	6	2.0	216	10	Q43430	Q43430 pisum sativ
704	6	2.0	199	10	Q80508	Q80508 arabidopsis	777	6	2.0	216	10	Q9ZSX5	Q9zxs5 zea mays (m
705	6	2.0	199	12	Q9QN25	Q9qn25 human immun	778	6	2.0	216	10	Q9SLG7	Q9slg7 arabidopsis
706	6	2.0	200	2	Q05796	Q05796 mycobacteri	779	6	2.0	217	4	Q9Z589	Q9z589 homo sapien
707	6	2.0	200	3	Q80003	Q80003 saccharomyc	780	6	2.0	218	2	P73776	P73776 synechocyst
708	6	2.0	200	3	Q13524	Q13524 schizophyll	781	6	2.0	218	2	Q9RS18	Q9rs18 deinococcus
709	6	2.0	200	10	Q9LWE5	Q9lwe5 oryza sativ	782	6	2.0	218	2	Q9PAD3	Q9pad3 xylella fas
710	6	2.0	200	11	Q88613	Q88613 mariones un	783	6	2.0	218	2	Q9KCB1	Q9kcb1 bacillus ha
711	6	2.0	200	12	Q70654	Q70654 porcine rep	784	6	2.0	218	4	Q9UDP2	Q9udp2 homo sapien
712	6	2.0	200	12	Q72701	Q72701 bunyavirus.	785	6	2.0	218	10	Q9SFK4	Q9sfk4 arabidopsis
713	6	2.0	200	12	Q72702	Q72702 bunyavirus.	786	6	2.0	218	10	Q9LY15	Q9ly15 arabidopsis
714	6	2.0	200	12	Q73535	Q73535 rift valley	787	6	2.0	218	12	Q66132	Q66132 cucumber mo
715	6	2.0	200	12	Q73536	Q73536 rift valley	788	6	2.0	218	12	Q89125	Q89125 cucumber mo
716	6	2.0	200	12	Q73537	Q73537 rift valley	789	6	2.0	218	12	Q9YJ53	Q9yjs3 cucumber mo
717	6	2.0	200	12	Q73538	Q73538 rift valley	790	6	2.0	220	5	Q20808	Q20808 caenorhabdi
718	6	2.0	200	12	Q73539	Q73539 rift valley	791	6	2.0	220	5	Q9W3E9	Q9w3e9 drosophila
719	6	2.0	200	12	Q73540	Q73540 rift valley	792	6	2.0	220	12	Q88461	Q88461 stealth vir
720	6	2.0	200	12	Q73541	Q73541 rift valley	793	6	2.0	221	13	Q90685	Q90685 gallus gall
721	6	2.0	200	12	Q73542	Q73542 rift valley	794	6	2.0	222	5	Q9XUI0	Q9xui0 caenorhabdi
722	6	2.0	200	12	Q73543	Q73543 rift valley	795	6	2.0	222	5	Q9W5M9	Q9w5m9 drosophila
723	6	2.0	200	12	Q73544	Q73544 rift valley	796	6	2.0	222	10	Q93106	Q93106 arabidopsis
724	6	2.0	200	12	Q73545	Q73545 rift valley	797	6	2.0	222	10	Q9ZVK9	Q9zvk9 arabidopsis
725	6	2.0	200	12	Q73546	Q73546 rift valley	798	6	2.0	222	10	Q9SLW7	Q9slw7 nicotiana t
726	6	2.0	200	12	Q73547	Q73547 rift valley	799	6	2.0	223	1	Q9YF37	Q9yfr37 aeropyrum p
727	6	2.0	200	12	Q73548	Q73548 rift valley	800	6	2.0	223	13	Q9YHV3	Q9yvh3 brachydanio
728	6	2.0	200	12	Q73549	Q73549 rift valley	801	6	2.0	224	2	Q9RFJ3	Q9rfj3 streptococc
729	6	2.0	200	12	Q89324	Q89324 porcine rep	802	6	2.0	224	5	Q9W3K8	Q9w3k8 drosophila
730	6	2.0	201	10	Q9XFY4	Q9xfy4 chlorella p	803	6	2.0	224	5	Q9V740	Q9v740 drosophila
731	6	2.0	202	2	Q07773	Q07773 mycobacteri	804	6	2.0	225	2	Q34343	Q34343 bacillus su
732	6	2.0	202	2	Q9S411	Q9s411 actinobacil	805	6	2.0	225	6	Q28271	Q28271 canis fami
733	6	2.0	202	5	Q91724	Q91724 drosophila	806	6	2.0	225	9	Q64371	Q64371 lactobacill
734	6	2.0	202	10	Q41111	Q41111 phaseolus v	807	6	2.0	225	10	Q9LHZ9	Q9lhz9 oryza sativ
735	6	2.0	202	10	Q9SW89	Q9sw89 prunus dulc	808	6	2.0	225	12	Q98708	Q98708 simian sarc
736	6	2.0	202	10	Q9SM68	Q9sm68 oryza sativ	809	6	2.0	226	2	Q47854	Q47854 erwina her
737	6	2.0	202	10	Q9LGP0	Q9lgp0 oryza sativ	810	6	2.0	226	2	P97177	P97177 rhodobacter
738	6	2.0	203	2	Q31942	Q31942 bacillus su	811	6	2.0	226	2	Q9KE61	Q9ke61 bacillus ha
739	6	2.0	203	2	Q87701	Q87701 bacillus me	812	6	2.0	226	10	Q9LLS1	Q9lls1 prunus dulc
740	6	2.0	203	2	Q9X875	Q9x875 streptomyc	813	6	2.0	227	4	Q9RYM6	Q9rym6 deinococcus
741	6	2.0	203	2	Q07795	Q07795 mycobacteri	814	6	2.0	227	4	Q94774	Q94774 homo sapien
742	6	2.0	203	9	Q64079	Q64079 bacterioph	815	6	2.0	227	10	Q9SSJ5	Q9ssj5 arabidopsis
743	6	2.0	203	10	Q9LDB6	Q9ldb6 catharanthu	816	6	2.0	227	10	Q9LJQ2	Q9ljq2 arabidopsis
744	6	2.0	204	2	Q9XDE2	Q9xde2 bruceella me	817	6	2.0	228	1	Q9LEY3	Q9ley3 aeropyrum p
745	6	2.0	204	2	Q9RKA5	Q9rka5 streptomyc	818	6	2.0	228	2	Q83764	Q83764 treponema p
746	6	2.0	204	5	Q9V4R8	Q9v4r8 drosophila	819	6	2.0	228	5	Q01329	Q01329 caenorhabdi
747	6	2.0	205	2	P95588	P95588 rhodobacter	820	6	2.0	228	5	Q96534	Q96534 strongyloce
748	6	2.0	205	10	Q43477	Q43477 hordeum vul	821	6	2.0	228	5	Q9XW11	Q9xw11 caenorhabdi
749	6	2.0	205	10	Q9SPA8	Q9spa8 hordeum vul	822	6	2.0	228	5	Q9NMU0	Q9nmu0 leishmania

823	6	2.0	228	12	Q69078	O69078 human herpe	896	6	2.0	243	10	Q9SP35	Q9SP35 arabidopsis
824	6	2.0	229	2	O67801	O67801 aquifex aeo	897	6	2.0	243	10	Q9LKA1	Q9LKA1 arabidopsis
825	6	2.0	229	2	Q843A4	Q843A4 streptococc	898	6	2.0	243	12	Q65818	Q65818 bovine herp
826	6	2.0	229	5	Q9W208	Q9W208 drosophila	899	6	2.0	244	2	Q45845	Q45845 clostridium
827	6	2.0	229	5	Q9N399	Q9N399 caenorhabdi	900	6	2.0	244	5	Q18352	Q18352 echinococcu
828	6	2.0	229	10	Q41723	Q41723 zinnia eleg	901	6	2.0	244	5	Q24902	Q24902 echinococcu
829	6	2.0	230	2	Q9S0A1	Q9S0A1 borrelia bu	902	6	2.0	244	5	Q9UAH0	Q9UAH0 tetrahymena
830	6	2.0	230	2	Q9R202	Q9R202 borrelia bu	903	6	2.0	244	5	Q9UA08	Q9UA08 echinococcu
831	6	2.0	230	2	Q9L911	Q9L911 salmonella	904	6	2.0	244	10	Q9SHD7	Q9SHD7 arabidopsis
832	6	2.0	230	5	Q18269	Q18269 caenorhabdi	905	6	2.0	245	5	Q9UAJ6	Q9UAJ6 plasmodium
833	6	2.0	230	10	Q9SCA7	Q9SCA7 lycopersico	906	6	2.0	245	13	Q57469	Q57469 xenopus lae
834	6	2.0	230	11	Q9QZL9	Q9QZL9 mus musculu	907	6	2.0	246	2	Q9XG66	Q9XG66 streptococc
835	6	2.0	231	2	Q9PIQ8	Q9PIQ8 campylobact	908	6	2.0	246	2	Q9S2T2	Q9S2T2 streptomyce
836	6	2.0	231	5	Q24973	Q24973 giardia lam	909	6	2.0	246	4	Q9UG67	Q9UG67 homo sapien
837	6	2.0	231	10	Q41579	Q41579 triticum ae	910	6	2.0	246	5	Q9VVA5	Q9VVA5 drosophila
838	6	2.0	232	2	Q48821	Q48821 lactobacill	911	6	2.0	246	11	Q70455	Q70455 mus musculu
839	6	2.0	232	2	Q9LIM7	Q9LIM7 streptomyce	912	6	2.0	246	11	O88301	O88301 mus musculu
840	6	2.0	232	5	Q22473	Q22473 caenorhabdi	913	6	2.0	247	2	P71851	P71851 mycobacteri
841	6	2.0	232	5	Q9XXX3	Q9XXX3 plasmodium	914	6	2.0	247	5	Q9NDA2	Q9NDA2 trypanosoma
842	6	2.0	232	10	Q9ZTR8	Q9ZTR8 hordeum vul	915	6	2.0	247	10	Q9ZTR6	Q9ZTR6 hordeum vul
843	6	2.0	232	10	Q9MOM3	Q9MOM3 arabidopsis	916	6	2.0	247	11	O70457	O70457 mus musculu
844	6	2.0	233	2	Q06528	Q06528 lactobacill	917	6	2.0	248	2	O68731	O68731 yersinia pe
845	6	2.0	233	2	Q04138	Q04138 lactococcus	918	6	2.0	248	5	O96945	O96945 geodia cydo
846	6	2.0	233	5	P92170	P92170 plasmodium	919	6	2.0	248	5	O9UAJ2	Q9UAJ2 plasmodium
847	6	2.0	233	10	O64498	O64498 arabidopsis	920	6	2.0	248	6	O77642	O77642 ovis aries
848	6	2.0	234	1	Q9Y9J1	Q9Y9J1 aeropyrum p	921	6	2.0	248	8	O9XML4	Q9XML4 pisum sativ
849	6	2.0	234	5	O61132	O61132 plasmodium	922	6	2.0	248	10	O80892	O80892 arabidopsis
850	6	2.0	234	5	O9XYW7	Q9XYW7 entodinium	923	6	2.0	248	10	Q9SWP7	Q9SWP7 arabidopsis
851	6	2.0	234	12	Q9YQ01	Q9YQ01 infectious	924	6	2.0	248	10	O70456	O70456 mus musculu
852	6	2.0	234	12	Q9IWL5	Q9IWL5 infectious	925	6	2.0	249	2	O9X8D5	Q9X8D5 streptomyce
853	6	2.0	234	12	Q9IWL3	Q9IWL3 infectious	926	6	2.0	249	2	O9RYL4	Q9RYL4 deinococcus
854	6	2.0	234	12	Q9IWL1	Q9IWL1 infectious	927	6	2.0	249	5	O9XXY1	Q9XXY1 plasmodium
855	6	2.0	234	12	Q9IWK9	Q9IWK9 infectious	928	6	2.0	249	5	Q9VSA8	Q9VSA8 drosophila
856	6	2.0	234	12	Q9IWK7	Q9IWK7 infectious	929	6	2.0	249	5	Q9U491	Q9U491 schistosoma
857	6	2.0	234	12	Q9IWK5	Q9IWK5 infectious	930	6	2.0	249	5	O9XTB1	Q9XTB1 caenorhabdi
858	6	2.0	234	12	Q9IWK3	Q9IWK3 infectious	931	6	2.0	249	10	P93787	P93787 solanum tub
859	6	2.0	234	12	Q9IWK1	Q9IWK1 infectious	932	6	2.0	249	10	Q9SCM4	Q9SCM4 arabidopsis
860	6	2.0	234	12	Q9IWT9	Q9IWT9 infectious	933	6	2.0	249	10	O49550	O49550 arabidopsis
861	6	2.0	234	12	Q9IWT7	Q9IWT7 infectious	934	6	2.0	249	12	Q9WIW9	Q9WIW9 infectious
862	6	2.0	235	2	Q9XCK9	Q9XCK9 streptococc	935	6	2.0	250	2	O30982	O30982 rhizobium h
863	6	2.0	235	2	Q9RLC4	Q9RLC4 pseudomonas	936	6	2.0	250	2	O9Z6C4	Q9Z6C4 pseudomonas
864	6	2.0	235	2	Q9REZ9	Q9REZ9 pseudomonas	937	6	2.0	250	2	O9XAB0	Q9XAB0 streptomyce
865	6	2.0	235	2	Q9REZ8	Q9REZ8 pseudomonas	938	6	2.0	250	2	Q9KW98	Q9KW98 agrobacteri
866	6	2.0	235	2	Q9REM4	Q9REM4 pseudomonas	939	6	2.0	250	7	O97981	Q97981 sus scrofa
867	6	2.0	235	2	Q9REL5	Q9REL5 pseudomonas	940	6	2.0	250	10	Q9M093	Q9M093 arabidopsis
868	6	2.0	235	2	Q9REL4	Q9REL4 pseudomonas	941	6	2.0	251	2	P95151	P95151 mycobacteri
869	6	2.0	235	5	Q96535	Q96535 strongyloce	942	6	2.0	251	11	O54854	O54854 rattus norv
870	6	2.0	235	10	Q9M283	Q9M283 arabidopsis	943	6	2.0	252	2	Q9L9I9	Q9L9I9 salmonella
871	6	2.0	236	5	Q19805	Q19805 caenorhabdi	944	6	2.0	252	5	Q9UAJ7	Q9UAJ7 plasmodium
872	6	2.0	236	10	Q9SLC3	Q9SLC3 arabidopsis	945	6	2.0	252	5	Q9UAJ5	Q9UAJ5 plasmodium
873	6	2.0	236	10	Q9SK78	Q9SK78 arabidopsis	946	6	2.0	252	6	Q9TV88	Q9TV88 sus scrofa
874	6	2.0	236	13	Q9I869	Q9I869 brachydanio	947	6	2.0	252	6	Q9TV87	Q9TV87 sus scrofa
875	6	2.0	237	5	Q9XXX6	Q9XXX6 plasmodium	948	6	2.0	252	6	Q9TV86	Q9TV86 sus scrofa
876	6	2.0	238	2	Q9XCF0	Q9XCF0 streptomyce	949	6	2.0	252	10	Q9SRS7	Q9SRS7 arabidopsis
877	6	2.0	238	2	Q9JZZ1	Q9JZZ1 neisseria m	950	6	2.0	253	2	Q9WYH9	Q9WYH9 thermotoga
878	6	2.0	238	3	O13547	O13547 saccharomyc	951	6	2.0	253	2	Q9S492	Q9S492 streptococc
879	6	2.0	238	4	Q99581	Q99581 homo sapien	952	6	2.0	253	2	Q9L018	Q9L018 streptomyce
880	6	2.0	238	5	O44958	O44958 caenorhabdi	953	6	2.0	253	5	O18298	O18298 caenorhabdi
881	6	2.0	238	5	Q9UAJ3	Q9UAJ3 plasmodium	954	6	2.0	253	5	Q9VBD2	Q9VBD2 drosophila
882	6	2.0	238	5	Q9UAJ1	Q9UAJ1 plasmodium	955	6	2.0	253	5	Q9UAJ4	Q9UAJ4 plasmodium
883	6	2.0	238	10	Q9LGH6	Q9LGH6 oryza sativ	956	6	2.0	253	9	P79670	P79670 bacterioph
884	6	2.0	239	2	Q9RTG3	Q9RTG3 deinococcus	957	6	2.0	253	10	Q9T007	Q9T007 arabidopsis
885	6	2.0	239	2	Q9KYG0	Q9KYG0 streptomyce	958	6	2.0	254	2	Q9KGH5	Q9KGH5 bacillus ha
886	6	2.0	239	5	O45375	O45375 caenorhabdi	959	6	2.0	254	5	Q9U8K4	Q9U8K4 schistosoma
887	6	2.0	240	2	O67575	O67575 aquifex aeo	960	6	2.0	254	10	P93785	P93785 solanum tub
888	6	2.0	240	5	Q9V8M0	Q9V8M0 drosophila	961	6	2.0	254	10	Q9LME5	Q9LME5 arabidopsis
889	6	2.0	241	5	P90579	P90579 plasmodium	962	6	2.0	254	13	Q91207	Q91207 oncorhynch
890	6	2.0	241	5	Q9XW21	Q9XW21 caenorhabdi	963	6	2.0	255	2	O84643	O84643 chlamydia t
891	6	2.0	241	10	Q9S9Z8	Q9S9Z8 arabidopsis	964	6	2.0	255	2	Q9R9L8	Q9R9L8 rhizobium m
892	6	2.0	242	1	Q9YAC1	Q9YAC1 aeropyrum p	965	6	2.0	255	4	Q9UGN7	Q9UGN7 homo sapien
893	6	2.0	243	5	P90527	P90527 dictyosteli	966	6	2.0	255	4	O82352	O82352 arabidopsis
894	6	2.0	243	5	P92169	P92169 plasmodium	967	6	2.0	255	12	P88301	P88301 human immun
895	6	2.0	243	10	Q9ZUS5	Q9ZUS5 arabidopsis	968	6	2.0	255	12	P90108	P90108 human immun

969 6 2.0 255 13 057468 057468 xenopus lae
 970 6 2.0 256 2 084256 084256 chlamydia t
 971 6 2.0 256 2 09pdl13 09pdl13 xylella fas
 972 6 2.0 256 2 09pct1 09pct1 xylella fas
 973 6 2.0 256 2 09p9v6 09p9v6 rhodobacter
 974 6 2.0 256 2 09kwr8 09kwr8 rhodobacter
 975 6 2.0 256 2 09kac2 09kac2 bacillus ha
 976 6 2.0 256 3 09ur29 09ur29 lentinula e
 977 6 2.0 256 5 021665 021665 caenorhabdi
 978 6 2.0 256 5 09xxx4 09xxx4 plasmodium
 979 6 2.0 256 5 017978 017978 caenorhabdi
 980 6 2.0 256 5 09nks5 09nks5 ciona intes
 981 6 2.0 256 10 024222 024222 oryza sativ
 982 6 2.0 257 2 071568 071568 mycobacteri
 983 6 2.0 257 2 09lcs6 09lcs6 lactobacill
 984 6 2.0 257 5 016943 016943 caenorhabdi
 985 6 2.0 257 5 028653 028653 sarcophaga
 986 6 2.0 257 5 09xxx5 09xxx5 plasmodium
 987 6 2.0 257 5 09t250 09t250 caenorhabdi
 988 6 2.0 258 2 09rk32 09rk32 streptomyc
 989 6 2.0 258 2 09rds1 09rds1 streptomyc
 990 6 2.0 258 5 001551 001551 caenorhabdi
 991 6 2.0 258 10 039937 039937 helianthus
 992 6 2.0 258 10 092tc2 092tc2 arabidopsis
 993 6 2.0 258 12 055501 055501 onion yello
 994 6 2.0 258 12 055502 055502 onion yello
 995 6 2.0 258 12 012680 012680 infectious
 996 6 2.0 259 2 051977 051977 pseudomonas
 997 6 2.0 259 5 09nbv7 09nbv7 schistosoma
 998 6 2.0 259 10 049082 049082 fritillaria
 999 6 2.0 259 10 093786 093786 solanum tub
 1000 6 2.0 259 10 023966 023966 helianthus

ALIGNMENTS

RESULT 1
 095407 ID Q95407 PRELIMINARY; PRT; 300 AA.
 AC Q95407;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE DECOY RECEPTOR 3 (M68) (M68C) (M68E).
 GN DCR3 OR TR6.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99087326; PubMed=9872321;
 RA Pitti R.M., Masters S.A., Lawrence D.A., Roy M., Kischkel F.C.,
 Dowd P., Huang A., Donahue C.J., Sherwood S.W., Baldwin D.T.,
 Godowski P.J., Wood W.I., Gurney A.L., Hillan K.J., Cohen R.L.,
 Goddard A.D., Botstein D., Ashkenazi A.;
 RA "Genomic amplification of a decoy receptor for Fas ligand in lung and
 RT colon cancer".
 RL Nature 396:699-703(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BLOOD;
 RX MEDLINE=99253915; PubMed=10318773;
 RA Yu K.Y., Kwon B., Ni J., Zhai Y., Ebner R., Kwon B.S.;
 RA "A newly identified member of tumor necrosis factor receptor
 RT superfamily (TR6) suppresses LfH1-mediated apoptosis".;
 RL J. Biol. Chem. 274:13733-13736(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PANCREAS;
 RX MEDLINE=20122600; PubMed=10655513;
 RA Bai C., Connolly B., Metzker M.L., Hilliard C.A., Liu X., Sandig V.,

RA Soderman A., Galloway S.M., Liu Q., Austin C.P., Caskey C.T.;
 RT "Overexpression of M68/Dcr3 in human gastrointestinal tract tumors
 RT independent of gene amplification and its location in a four-gene
 RT cluster".
 RL Proc. Natl. Acad. Sci. U.S.A. 97:1230-1235(2000).
 DR EMBL; AF104419; AAD03056.1; -;
 DR EMBL; AF134240; AAD29688.1; -;
 DR EMBL; AF217796; AAF35244.1; -;
 DR EMBL; AF217793; AAF33685.1; -;
 DR EMBL; AF217794; AAF33686.1; -;
 DR HSSP; P25942; 1CDF.
 DR INTERPRO; IPR000561; -;
 DR INTERPRO; IPR001368; -;
 DR PFAM; PF00020; TNFR_C6; 4.
 DR PROSITE; PS00652; TNFR_NGFR_1; UNKNOWN_1.
 DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
 DR PROSITE; PS50050; TNFR_NGFR_2; 2.
 DR PRODOM; PD000771; -; 1.
 KW Receptor.
 SQ SEQUENCE 300 AA; 32679 MW; F90AEE33718449AF CRC64;
 Query Match 100.0%; Score 300; DB 4; Length 300;
 Best Local Similarity 100.0%; Pred. No. 1.le-279;
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRALGEGLSLCLVLALPALLPVPVAVGVAETPYPMWDAETGERLVCAQCPPTGVQR 60
 DB 1 MRALGEGLSLCLVLALPALLPVPVAVGVAETPYPMWDAETGERLVCAQCPPTGVQR 60
 QY 61 PCRRDSPTTCGPPRHVYTFQWNYLERCYCNVLCGEREEERACHATHNRACRGTGFF 120
 DB 61 PCRRDSPTTCGPPRHVYTFQWNYLERCYCNVLCGEREEERACHATHNRACRGTGFF 120
 QY 121 AHAGCLEHASCPPGAGVIAPCTPSONTCOCPCPCTFSSSSSQOCPHNRCTALGLA 180
 DB 121 AHAGCLEHASCPPGAGVIAPCTPSONTCOCPCPCTFSSSSSQOCPHNRCTALGLA 180
 QY 181 LNVPGSSSHDTLCTCTGTFPLSTRVPGAECEERAVIDFVAFQDISIKRLQLQALEAPE 240
 DB 181 LNVPGSSSHDTLCTCTGTFPLSTRVPGAECEERAVIDFVAFQDISIKRLQLQALEAPE 240
 QY 241 GWGPTPRAGRAALQLKRLRLTELLGAQDQALLVRLQLALRVARMPGLERSVRERFLPVH 300
 DB 241 GWGPTPRAGRAALQLKRLRLTELLGAQDQALLVRLQLALRVARMPGLERSVRERFLPVH 300
 RESULT 2
 09N092 ID Q9N092 PRELIMINARY; PRT; 430 AA.
 AC Q9N092;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE UNNAMED PROTEIN PRODUCT.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Cercopitheidae;
 OC Cercopitheidae; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
 Suzuki Y., Sugano S., Hashimoto K.;
 RA "Isolation of full-length cDNA clones from macaque brain cDNA
 RT libraries".
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB046039; BAB01621.1; -;
 SQ SEQUENCE 430 AA; 45850 MW; BA8DE92593E1E859 CRC64;
 Query Match 3.0%; Score 9; DB 6; Length 430;
 Best Local Similarity 100.0%; Pred. No. 2.5;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 PCPPGTFSA 160
|||||

Db 50 PCPPGTFSA 58

RESULT 3

ID Q9SH82 PRELIMINARY; PRT; 561 AA.
AC Q9SH82;
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE PUTATIVE NA+-DEPENDENT INORGANIC PHOSPHATE COTRANSPORTER.
GN ATG38060.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.W., Somerville C.R.,
RA Copenhaver G.P., Preuss C.M., Venter J.C.;
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome II of Arabidopsis thaliana.";
RL Nature 402:761-768(1999).
DR EMBL: AC007661; AAD32766.1; -;
SQ SEQUENCE 561 AA; 61232 MW; EEFB0BF3127E7680 CRC64;

Query Match 3.0%; Score 9; DB 10; Length 561;

Best Local Similarity 100.0%; Pred. No. 3.2;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 GPGLSLCL 14

Db 405 GPGLSLCL 413

RESULT 4

ID O99888 PRELIMINARY; PRT; 54 AA.
AC O99888;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE LARGE T AND SMALL T ANTIGENS (FRAGMENT).
OS Budgerigar fledgling disease virus (BFDV).
OC Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Polyomavirus.
OX NCBI_TaxID=10625;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MCFL97;
RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF054402; AAC33626.1; -;
DR HSPF: P25685; 1HDJ.
DR INTERPRO: IPR001623; -;
DR PFAM: PF00226; DnaJ; 1.
DR PROSITE: PS50076; DNAJ_2; 1.
FT NON_TER 54
SQ SEQUENCE 54 AA; 6077 MW; 5AF094925DE8C997 CRC64;

Query Match

2.7%; Score 8; DB 12; Length 54;

Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 259 RRLTELLG 266

Db 5 RRLTELLG 12

RESULT 5

ID O99889 PRELIMINARY; PRT; 54 AA.
AC O99889;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE LARGE T AND SMALL T ANTIGENS (FRAGMENT).
OS Budgerigar fledgling disease virus (BFDV).
OC Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Polyomavirus.
OX NCBI_TaxID=10625;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LBMI92;
RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF054403; AAC33627.1; -;
DR INTERPRO: IPR001623; -;
DR PFAM: PF00226; DnaJ; 1.
DR PROSITE: PS50076; DNAJ_2; 1.
FT NON_TER 54
SQ SEQUENCE 54 AA; 6077 MW; 5AF094925DE8C997 CRC64;

Query Match

2.7%; Score 8; DB 12; Length 54;

Best Local Similarity 100.0%; Pred. No. 4.1;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 259 RRLTELLG 266

Db 5 RRLTELLG 12

RESULT 6

ID O99890 PRELIMINARY; PRT; 54 AA.
AC O99890;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE LARGE T AND SMALL T ANTIGENS (FRAGMENT).
OS Budgerigar fledgling disease virus (BFDV).
OC Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Polyomavirus.
OX NCBI_TaxID=10625;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GCAZ92;
RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF054404; AAC33628.1; -;
DR INTERPRO: IPR001623; -;
DR PFAM: PF00226; DnaJ; 1.
DR PROSITE: PS50076; DNAJ_2; 1.
FT NON_TER 54
SQ SEQUENCE 54 AA; 6077 MW; 5AF094925DE8C997 CRC64;

Query Match

2.7%; Score 8; DB 12; Length 54;

Best Local Similarity 100.0%; Pred. No. 4.1;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 259 RRLTELLG 266

|||||

```

Db      5 RRLTELLG 12

RESULT  7
ID      089891      PRELIMINARY;      PRT;      54 AA.
AC      089891;
DT      01-NOV-1998 (TREMBLrel. 08, Created)
DT      01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT      01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE      LARGE T AND SMALL T ANTIGENS (FRAGMENT).
OS      Budgerigar fledgling disease virus (BFDV).
OC      Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Polyomavirus.
OX      NCBI_TaxID=10625;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ECTX91;
RA      Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
RT      "Genetic Diversity in 20 Variants of the Avian Polyomavirus.";
RL      Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AF054405; AAC33629.1; -
DR      INTERPRO: IPR001623; -
DR      PFAM: PF00226; DnaJ; 1.
DR      PROSITE: PS50076; DNAJ_2; 1.
FT      NON_TER      54      54
SQ      SEQUENCE      54 AA; 6077 MW; 5AF094925DE8C997 CRC64;

Query Match      2.7%; Score 8; DB 12; Length 54;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

QY      259 RRLTELLG 266
Db      5 RRLTELLG 12
|||||||

RESULT  8
ID      089892      PRELIMINARY;      PRT;      54 AA.
AC      089892;
DT      01-NOV-1998 (TREMBLrel. 08, Created)
DT      01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT      01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE      LARGE T AND SMALL T ANTIGENS (FRAGMENT).
OS      Budgerigar fledgling disease virus (BFDV).
OC      Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Polyomavirus.
OX      NCBI_TaxID=10625;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=LB85;
RA      Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
RT      "Genetic Diversity in 20 Variants of the Avian Polyomavirus.";
RL      Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AF054406; AAC33630.1; -
DR      INTERPRO: IPR001623; -
DR      PFAM: PF00226; DnaJ; 1.
DR      PROSITE: PS50076; DNAJ_2; 1.
FT      NON_TER      54      54
SQ      SEQUENCE      54 AA; 6077 MW; 5AF094925DE8C997 CRC64;

Query Match      2.7%; Score 8; DB 12; Length 54;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

QY      259 RRLTELLG 266
Db      5 RRLTELLG 12
|||||||

RESULT  9
ID      089893      PRELIMINARY;      PRT;      54 AA.
AC      089893;
DT      01-NOV-1998 (TREMBLrel. 08, Created)
DT      01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT      01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE      LARGE T AND SMALL T ANTIGENS (FRAGMENT).
OS      Budgerigar fledgling disease virus (BFDV).
OC      Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Polyomavirus.
OX      NCBI_TaxID=10625;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=YCFL87;
RA      Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
RT      "Genetic Diversity in 20 Variants of the Avian Polyomavirus.";
RL      Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AF054407; AAC33631.1; -
DR      INTERPRO: IPR001623; -
DR      PFAM: PF00226; DnaJ; 1.
DR      PROSITE: PS50076; DNAJ_2; 1.
FT      NON_TER      54      54
SQ      SEQUENCE      54 AA; 6077 MW; 5AF094925DE8C997 CRC64;

Query Match      2.7%; Score 8; DB 12; Length 54;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

QY      259 RRLTELLG 266
Db      5 RRLTELLG 12
|||||||

RESULT 10
ID      089894      PRELIMINARY;      PRT;      54 AA.
AC      089894;
DT      01-NOV-1998 (TREMBLrel. 08, Created)
DT      01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT      01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE      LARGE T AND SMALL T ANTIGENS (FRAGMENT).
OS      Budgerigar fledgling disease virus (BFDV).
OC      Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Polyomavirus.
OX      NCBI_TaxID=10625;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=SCIL88;
RA      Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
RT      "Genetic Diversity in 20 Variants of the Avian Polyomavirus.";
RL      Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AF054408; AAC33632.1; -
DR      INTERPRO: IPR001623; -
DR      PFAM: PF00226; DnaJ; 1.
DR      PROSITE: PS50076; DNAJ_2; 1.
FT      NON_TER      54      54
SQ      SEQUENCE      54 AA; 6077 MW; 5AF094925DE8C997 CRC64;

Query Match      2.7%; Score 8; DB 12; Length 54;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

QY      259 RRLTELLG 266
Db      5 RRLTELLG 12
|||||||

RESULT 11
ID      089895      PRELIMINARY;      PRT;      54 AA.
AC      089895;
DT      01-NOV-1998 (TREMBLrel. 08, Created)
DT      01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT      01-OCT-2000 (TREMBLrel. 15, Last annotation update)

```

```
DE LARGE T AND SMALL T ANTIGENS (FRAGMENT).
OS Budgerigar fledgling disease virus (BFDV).
OC Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Polyomavirus.
OX NCBI_TaxID=10625;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BDTX89;
RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF054410; AAC33633.1; -.
DR INTERPRO; IPR001623; -.
DR PFAM; PF00226; DnaJ; 1.
DR PROSITE; PS50076; DNAJ_2; 1.
DR NON_TER 54
FT SEQUENCE 54 AA; 6077 MW; 5AF094925DE8C997 CRC64;
SQ

Query Match      2.7%; Score 8; DB 12; Length 54;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 259 RRLTELLG 266
Db 5 RRLTELLG 12

RESULT 12
O89896 PRELIMINARY; PRT; 54 AA.
AC O89896;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE LARGE T AND SMALL T ANTIGENS (FRAGMENT).
OS Budgerigar fledgling disease virus (BFDV).
OC Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Polyomavirus.
OX NCBI_TaxID=10625;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BDTX89;
RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF054410; AAC33633.1; -.
DR INTERPRO; IPR001623; -.
DR PFAM; PF00226; DnaJ; 1.
DR PROSITE; PS50076; DNAJ_2; 1.
DR NON_TER 54
FT SEQUENCE 54 AA; 6077 MW; 5AF094925DE8C997 CRC64;
SQ

Query Match      2.7%; Score 8; DB 12; Length 54;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 259 RRLTELLG 266
Db 5 RRLTELLG 12

RESULT 13
O89897 PRELIMINARY; PRT; 54 AA.
AC O89897;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE LARGE T AND SMALL T ANTIGENS (FRAGMENT).
OS Budgerigar fledgling disease virus (BFDV).
OC Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Polyomavirus.
OX NCBI_TaxID=10625;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=BDTX89;
RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF054411; AAC33635.1; -.
DR INTERPRO; IPR001623; -.
DR PFAM; PF00226; DnaJ; 1.
DR PROSITE; PS50076; DNAJ_2; 1.
DR NON_TER 54
FT SEQUENCE 54 AA; 6077 MW; 5AF094925DE8C997 CRC64;
SQ

Query Match      2.7%; Score 8; DB 12; Length 54;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 259 RRLTELLG 266
Db 5 RRLTELLG 12

RESULT 14
O89898 PRELIMINARY; PRT; 54 AA.
AC O89898;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE LARGE T AND SMALL T ANTIGENS (FRAGMENT).
OS Budgerigar fledgling disease virus (BFDV).
OC Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Polyomavirus.
OX NCBI_TaxID=10625;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BDGAB1-A;
RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF054412; AAC33636.1; -.
DR INTERPRO; IPR001623; -.
DR PFAM; PF00226; DnaJ; 1.
DR PROSITE; PS50076; DNAJ_2; 1.
DR NON_TER 54
FT SEQUENCE 54 AA; 6077 MW; 5AF094925DE8C997 CRC64;
SQ

Query Match      2.7%; Score 8; DB 12; Length 54;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 259 RRLTELLG 266
Db 5 RRLTELLG 12

RESULT 15
O89899 PRELIMINARY; PRT; 54 AA.
AC O89899;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE LARGE T AND SMALL T ANTIGENS (FRAGMENT).
OS Budgerigar fledgling disease virus (BFDV).
OC Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Polyomavirus.
OX NCBI_TaxID=10625;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BDGAB1-B;
RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
```


Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 259 RRLTELLG 266
 |||||
 Db 5 RRLTELLG 12

RESULT 20

O89904 PRELIMINARY; PRT; 54 AA.
 AC O89904;
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE LARGE T AND SMALL T ANTIGENS (FRAGMENT).
 OS Budgerigar fledgling disease virus (BFDV).
 OC Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Polyomavirus.
 OX NCBI_TaxID=10625;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BGTX93;
 RA Phalen D.N., Wilson V.G., Gaskin J.M., Derr J.N., Graham D.L.;
 RT "Genetic Diversity in 20 Variants of the Avian Polyomavirus";
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF054418; AAC33642.1; -.
 DR INTERPRO; IPR001623; -.
 DR PFAM; PF00226; DnaJ.1.
 DR PROSITE; PS50076; DnaJ.2; 1.
 FT NON_TER 54
 SQ SEQUENCE 54 AA; 6077 MW; 5AF094925DB8C997 CRC64;

Query Match 2.7%; Score 8; DB 12; Length 54;

Best Local Similarity 100.0%; Pred. No. 4.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 259 RRLTELLG 266
 |||||
 Db 5 RRLTELLG 12

RESULT 21

ID Q9TTA6 PRELIMINARY; PRT; 129 AA.
 AC Q9TTA6;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE PREPRO-OREXIN PRECURSOR (FRAGMENT).
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Malek M., Marklund S., Rothschild M.F.;
 RT "Linkage and physical mapping of the porcine prepro-orexin gene";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF169352; AAF24216.1; -.
 DR INTERPRO; IPR001704; -.
 DR PFAM; PF02072; Orexin; 1.
 DR PRINTS; PR01091; OREXINPP.
 KW Signal.
 FT NON_TER 1 1
 FT SIGNAL <1 5 POTENTIAL.
 SQ SEQUENCE 129 AA; 13211 MW; 0F5B9BE9C2D9AB7B CRC64;

Query Match 2.7%; Score 8; DB 6; Length 129;

Best Local Similarity 100.0%; Pred. No. 8.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 228 RLQRLQA 235

Db 77 RLQRLQA 84
 |||||
 |||||

RESULT 22

Q9WG04 PRELIMINARY; PRT; 145 AA.
 ID Q9WG04;
 AC Q9WG04;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE SMALL T ANTIGEN.
 OS Budgerigar fledgling disease virus (BFDV), and
 OS Budgerigar fledgling disease virus - 5.
 OC Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Polyomavirus.
 OX NCBI_TaxID=10625, 133791;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Budgerigar fledgling disease virus;
 RA Lafferty S.L., Fudge A.M., Schmidt R.E., Wilson V.G., Phalen D.N.;
 RT "Avian polyomavirus infection and disease in a green aracaris
 (Pteroglossus viridis).";
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Budgerigar fledgling disease virus - 5;
 RX MEDLINE-88265888; PubMed-2838972;
 RA Rott O., Kroger M., Muller H., Hobom G.;
 RT "The genome of budgerigar fledgling disease virus, an avian
 polyomavirus.";
 RL Virology 165:74-86(1988).
 [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Budgerigar fledgling disease virus - 5;
 RX MEDLINE-93155646; PubMed-6381462;
 RA Stoll R., Luo D., Kouwenhoven B., Hobom G., Muller H.;
 RT "Molecular and biological characteristics of avian polyomaviruses:
 isolates from different species of birds indicate that avian
 polyomaviruses form a distinct subgenus within the polyomavirus
 genus.";
 RL J. Gen. Virol. 74:229-237(1993).
 [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Budgerigar fledgling disease virus - 5;
 RX MEDLINE-95146945; PubMed-7844527;
 RA Luo D., Muller H., Tang X.B., Hobom G.;
 RT "Early and late pre-mRNA processing of budgerigar fledgling disease
 virus 1: identification of viral RNA 5' and 3' ends and internal
 splice junctions.";
 RL J. Gen. Virol. 76:161-166(1995).
 [5]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Budgerigar fledgling disease virus - 5;
 RX MEDLINE-98411646; PubMed-9739329;
 RA Johne R., Muller H.;
 RT "Avian polyomavirus in wild birds: genome analysis of isolates from
 Falconiformes and Psittaciformes.";
 RL Arch. Virol. 143:1501-1512(1998).
 [6]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Budgerigar fledgling disease virus - 5;
 RA Johne R., Jungmann A., Muller H.;
 RT "Agnoptrotein 1a and agnoptrotein 1b of avian polyomavirus are apoptotic
 inducers.";
 RL J. Gen. Virol. 0:0-0(2000).
 DR EMBL; AF118150; AAD30962.1; -.
 DR EMBL; AF241170; AAF97894.1; -.
 DR INTERPRO; IPR001623; -.
 DR PFAM; PF00226; DnaJ.1.
 DR PROSITE; PS50076; DnaJ.2; 1.
 SQ SEQUENCE 145 AA; 16942 MW; CD4AC88E3620D06B CRC64;

```

Query Match          2.7%; Score 8; DB 12; Length 145;
Best Local Similarity 100.0%; Pred. No. 9.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 259 RRLTELLG 266
DB 5 RRLTELLG 12

RESULT 23
Q9IBL0 PRELIMINARY; PRT; 145 AA.
ID Q9IBL0
AC Q9IBL0
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE SMALL T ANTIGEN.
OS Budgerigar fledgling disease virus - 4.
OC Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Polyomavirus.
OX NCBI_TaxID=133790;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88265888; PubMed=2838972;
RA Rott O., Kroger M., Muller H., Hobom G.;
RT "The genome of budgerigar fledgling disease virus, an avian
RT polyomavirus.";
RL Virology 165:74-86(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93155648; PubMed=8381462;
RA Stoll R., Luo D., Kouwenhoven B., Hobom G., Muller H.;
RT "Molecular and biological characteristics of avian polyomaviruses:
RT isolates from different species of birds indicate that avian
RT polyomaviruses form a distinct subgenus within the polyomavirus
RT genus.";
RL J. Gen. Virol. 74:229-237(1993).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=95146945; PubMed=7844527;
RA Luo D., Muller H., Tang X.B., Hobom G.;
RT "Early and late pre-mRNA processing of budgerigar fledgling disease
RT virus 1: identification of viral RNA 5' and 3' ends and internal
RT splice junctions.";
RL J. Gen. Virol. 76:161-166(1995).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=98411646; PubMed=9739329;
RA John R., Muller H.;
RT "Avian polyomavirus in wild birds: genome analysis of isolates from
RT Falconiformes and Psittaciformes.";
RL Arch. Virol. 143:1501-1512(1998).
RN [5]
RP SEQUENCE FROM N.A.
RA John R., Jungmann A., Muller H.;
RT "Agno protein 1a and agno protein 1b of avian polyomavirus are apoptotic
RT inducers.";
RL J. Gen. Virol. 0:0-0(2000).
DR ENBL: AF241169; AAF97887.1;
SQ SEQUENCE 145 AA; 16908 MW; 9DE0C88E3620D066 CRC64;

Query Match          2.7%; Score 8; DB 12; Length 145;
Best Local Similarity 100.0%; Pred. No. 9.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 259 RRLTELLG 266
DB 5 RRLTELLG 12

RESULT 24
O65446 PRELIMINARY; PRT; 181 AA.
ID O65446

```

```

O65446;
AC 01-AUG-1998 (TRENBLrel. 07, Created)
DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE HYPOTHETICAL 20.2 KDA PROTEIN.
GN FLN20.70 OR AT4G21970.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Wedler H., Wambutt R., Bancroft I., Mewes H.W., Mayer K.,
RA Schueller C.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Wedler H., Wambutt R., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project.
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL022140; CAA18100.1;
DR EMBL; AL161556; CAB79152.1;
DR MENDEL; 29060; Arath; 3376; 29060.
KW Hypothetical protein.
SQ SEQUENCE 181 AA; 20207 MW; 63BC254EEC39BF74 CRC64;

Query Match          2.7%; Score 8; DB 10; Length 181;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 158 FSASSSS 165
DB 35 FSASSSS 42

RESULT 25
Q9X6H9 PRELIMINARY; PRT; 206 AA.
ID Q9X6H9
AC Q9X6H9
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
DE SERUM OPACITY FACTOR PRECURSOR (FRAGMENT).
GN SOF.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SS985 M TYPE 63;
RA Beall B., Gherardi G., Li Z.;
RT "The relation of Streptococcus pyogenes sof and emm gene sequence
RT types to genetically distinct strain sets.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF138806; AAD31499.1;
KW Signal.
FT NON_TER 1 1
FT SIGNAL <1 12 POTENTIAL.
FT CHAIN 13 >206 SERUM OPACITY FACTOR.
FT NON_TER 206 206
FT NON_TER 206 206
SQ SEQUENCE 206 AA; 20802 MW; 3530DCEEF146ADF2B CRC64;

Query Match          2.7%; Score 8; DB 2; Length 206;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 159 SASSSSSE 166
 Db 87 SASSSSSE 94

RESULT 26

ID O49719 PRELIMINARY; PRT; 241 AA.
 AC O49719;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE HYPOTHETICAL 26.7 KDA PROTEIN.
 GN T805.180.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Wedler H., Wambutt R., Bancroft I., Mewes H.W., Mayer K.,
 RA Schueller C.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL021890; CAA17164.1; -;
 DR MENDEL; 28258; Arath; 3376; 28258.
 DR INTERPRO; IPR000215; -;
 DR PROSITE; PS00284; SERPIN; UNKNOWN_1.
 KW Hypothetical protein.
 SQ SEQUENCE 241 AA; 26705 MW; 533259379F5A1BC6 CRC64;

Query Match 2.7%; Score 8; DB 10; Length 241;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 158 FSASSSSS 165
 Db 35 FSASSSSS 42

RESULT 27

ID Q9XXL8 PRELIMINARY; PRT; 341 AA.
 AC Q9XXL8;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE ZK1025.9 PROTEIN.
 GN ZK1025.9.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lennard N.;
 RX MEDLINE=94150718; PubMed=7906398;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.";
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R.,
 RA Skidmore N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,

RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans";
 RL Nature 366:32-38(1994).
 DR EMBL; AL022288; CAA18368.1; -;
 DR HSSP; P19793; 2NLL.
 DR INTERPRO; IPR001628; -;
 DR PFAM; PF00105; zf-C4; 1.
 DR PRINTS; PR00047; STROIDFINGER.
 SQ SEQUENCE 341 AA; 38835 MW; 258833575F585D30 CRC64;

Query Match 2.7%; Score 8; DB 5; Length 341;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 159 SASSSSSE 166
 Db 103 SASSSSSE 110

RESULT 28

ID Q9JP96 PRELIMINARY; PRT; 440 AA.
 AC Q9JP96;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE ORF440 PROTEIN.
 OS Rhodocyclus gelatinosus (Rhodopseudomonas gelatinosa).
 OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae;
 OC Rubrivivax.
 OX NCBI_TaxID=28068;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IL144;
 RA Nagashima K.V., Shimada K., Matsuura K.;
 RT "Phylogenetic analysis of photosynthetic genes of Rhodocyclus
 RT gelatinosus: Possibility of horizontal gene transfer in purple
 RT bacteria.";
 RL Photosyn. Res. 36:185-191(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IL144;
 RX MEDLINE=94132007; PubMed=8300574;
 RA Nagashima K.V., Matsuura K., Ohya S., Shimada K.;
 RT "Primary structure and transcription of genes encoding B870 and
 RT photosynthetic reaction center apoproteins from Rubrivivax
 RT gelatinosus.";
 RL J. Biol. Chem. 269:2477-2484(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IL144;
 RA Igarashi N., Shimada K., Matsuura K., Nagashima K.V.;
 RT "PHOTOSYNTHETIC GENE CLUSTER IN PURPLE BACTERIUM, RUBRIVIVAX
 RT GELATINOSUS.";
 RL (in) Garab G. (eds.);
 RL Photosynthesis. Mechanisms and Effects IV:2889-2892;
 RL Kluwer Academic Publishing (1998).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IL144;
 RA Menin L., Yoshida M., Jaquinod M., Nagashima K.V., Matsuura K.,
 RA Parot P., Vermeglio A.;
 RT "Dark aerobic growth conditions induce the synthesis of a high
 RT midpoint potential cytochrome c8 in the photosynthetic bacterium
 RT Rubrivivax gelatinosus.";
 RL Biochemistry 0:0-0(1999).
 DR EMBL; AB034704; BAA94065.1; -;
 SQ SEQUENCE 440 AA; 45186 MW; 53238733B4F92A34 CRC64;

Query Match 2.7%; Score 8; DB 2; Length 440;

Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LALPALLP 23
| | | | | | | |
DB 42 LALPALLP 49

RESULT 29
QY577 PRELIMINARY: PRT; 477 AA.
AC QY577;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE RING FINGER PROTEIN TERF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RX MEDLINE=99011410; PubMed=9792805;
RA Ogawa S., Goto W., Orimo A., Hosoi T., Ouchi Y., Muramatsu M.,
RA Inoue S.;
RT "Molecular cloning of a novel RING finger-B box-coiled coil (RBCC)
protein, terf, expressed in the testis."
RL Biochem. Biophys. Res. Commun. 251:515-519(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RA Ogawa S., Goto W., Orimo A., Hosoi T., Ouchi Y., Muramatsu M.,
RA Inoue S.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS A C3HC4-CLASS ZINC FINGER.
DR EMBL; AF156271; RAD40286.1; -.
DR INTERPRO; IPR000107; -.
DR INTERPRO; IPR000315; -.
DR INTERPRO; IPR001841; -.
DR INTERPRO; IPR001870; -.
DR INTERPRO; IPR002991; -.
DR PFAM; PF00097; zf-C3HC4; 1.
DR PFAM; PF00622; SPRY; 1.
DR PFAM; PF00643; zf-B_box; 1.
DR PROSITE; PS00518; ZINC_FINGER_C3HC4; 1.
KW Zinc-finger.
SQ SEQUENCE 477 AA; 54418 MW; ECA4010661ADD28A CRC64;

Query Match 2.7%; Score 8; DB 4; Length 477;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 QRLQALE 237
| | | | | | | |
DB 201 QRLQALE 208

RESULT 30
Q42582 PRELIMINARY: PRT; 531 AA.
AC Q42582;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE 1-PHOSPHATIDYLINOSITOL-4,5-BISPHOSPHATE PHOSPHODIESTERASE
DE (EC 3.1.4.11) (TRIPHOSPHOINOSITIDE PHOSPHODIESTERASE).
GN PLC2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Kopka J., Willmitzer L., Mueller-Roeber B.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: 1-PHOSPHATIDYL-ID-MYO-INOSITOL 4,5-BISPHOSPHATE + H(2)O = D-MYO-INOSITOL 1,4,5-TRISPHOSPHATE + DIACYLGLYCEROL.
CC EMBL; X85973; CAA59962.1; -.
DR HSSP; P10688; 1DJH.
DR MENDEL; 6393; Arath; 1103; 6393.
DR INTERPRO; IPR000008; -.
DR INTERPRO; IPR000909; -.
DR INTERPRO; IPR001192; -.
DR INTERPRO; IPR001711; -.
DR PFAM; PF00168; C2; 1.
DR PFAM; PF00387; PI-PLC-Y; 1.
DR PFAM; PF00388; PI-PLC-X; 1.
DR PRINTS; PR00390; PHPLIPASEC.
DR PROSITE; PS50004; C2_DOMAIN_2; 1.
DR PROSITE; PS50007; PIPLC_X_DOMAIN; 1.
DR PROSITE; PS50008; PIPLC_Y_DOMAIN; 1.
DR PRODOM; PD001202; -; 1.
KW Hydrolase.
SQ SEQUENCE 531 AA; 61061 MW; 5EBCC5D0B45F93AA CRC64;

Query Match 2.7%; Score 8; DB 10; Length 531;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 259 RRLTELLG 266
| | | | | | | |
DB 38 RRLTELLG 45

RESULT 31
Q9STZ3 PRELIMINARY: PRT; 531 AA.
AC Q9STZ3;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE 1-PHOSPHATIDYLINOSITOL-4,5-BISPHOSPHATE PHOSPHODIESTERASE
DE (EC 3.1.4.11).
GN T2118.40.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Choisne N., Robert C., Brottier P., Wincker P., Cattolico L., Mayer K.F.X.,
RA Artiguenave F., Saurin W., Weissenbach J., Mewes H.W., Mayer K.F.X.,
RA Lemcke K., Schueller C., Quetier F., Salanoubat M.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL096860; CAB51201.1; -.
DR HSSP; P10688; 1DJH.
DR INTERPRO; IPR000008; -.
DR INTERPRO; IPR000909; -.
DR INTERPRO; IPR001192; -.
DR INTERPRO; IPR001711; -.
DR PFAM; PF00168; C2; 1.
DR PFAM; PF00387; PI-PLC-Y; 1.
DR PFAM; PF00388; PI-PLC-X; 1.
DR PRINTS; PR00390; PHPLIPASEC.
DR PROSITE; PS50004; C2_DOMAIN_2; 1.
DR PROSITE; PS50007; PIPLC_X_DOMAIN; 1.

DR PROSITE; PS50008; PIPLC_Y_DOMAIN; 1.
 KW Hydrolase.
 SQ SEQUENCE 531 AA; 61076 MW; BA4C2C960029771F CRC64;

Query Match 2.7%; Score 8; DB 10; Length 531;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 259 RRLTELLG 266
 |||||
 DB 38 RRLTELLG 45

RESULT 32

Q9WG03 ID Q9WG03 PRELIMINARY; PRT; 587 AA.

AC Q9WG03;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE LARGE T ANTIGEN.
 OS Budgerigar fledgling disease virus (BFDV).
 OC Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Polyomavirus.
 OX NCBI_TaxID=10625;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Laferriere S.L., Fudge A.M., Schmidt R.E., Wilson V.G., Phalen D.N.;
 RT "Avian polyomavirus infection and disease in a green arachnid
 (Pteroglossus viridis).";
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF118150; AAD30961.1; -;
 DR HSSP; P08622; 1BQZ.
 DR INTERPRO; IPR001623; -;
 DR PFAM; PF00226; DnaJ; 1.
 DR PROSITE; PS50076; DnaJ_2; 1.
 SQ SEQUENCE 587 AA; 67105 MW; 59CCBC02BBCC567 CRC64;

Query Match 2.7%; Score 8; DB 12; Length 587;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 259 RRLTELLG 266
 |||||
 DB 5 RRLTELLG 12

RESULT 33

Q9IBL1 ID Q9IBL1 PRELIMINARY; PRT; 599 AA.

AC Q9IBL1;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE LARGE T ANTIGEN.
 OS Budgerigar fledgling disease virus - 4.
 OC Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Polyomavirus.
 OX NCBI_TaxID=133790;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=88265888; PubMed=2838972;
 RA Rott O., Kroger M., Muller H., Hobom G.;
 RT "The genome of budgerigar fledgling disease virus, an avian
 polyomavirus.";
 RL Virology 165:74-86(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=93155648; PubMed=8381462;
 RA Stoll R., Luo D., Kouwenhoven B., Hobom G., Muller H.;
 RT "Molecular and biological characteristics of avian polyomaviruses:
 isolates from different species of birds indicate that avian
 polyomaviruses form a distinct subgenus within the polyomavirus
 genus.";
 RL Virology 165:74-86(1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=93155648; PubMed=8381462;
 RA Stoll R., Luo D., Kouwenhoven B., Hobom G., Muller H.;
 RT "Molecular and biological characteristics of avian polyomaviruses:
 isolates from different species of birds indicate that avian
 polyomaviruses form a distinct subgenus within the polyomavirus
 genus.";
 RL Virology 165:74-86(1988).
 RN [4]
 RP SEQUENCE FROM N.A.

RT J. Gen. Virol. 74:229-237(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=95146945; PubMed=7844527;
 RA Luo D., Muller H., Tang X.B., Hobom G.;
 RT "Early and late pre-mRNA processing of budgerigar fledgling disease
 virus 1: identification of viral RNA 5' and 3' ends and internal
 splice junctions.";
 RL J. Gen. Virol. 76:161-166(1995).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=98411646; PubMed=9739329;
 RA John R., Muller H.;
 RT "Avian polyomavirus in wild birds: genome analysis of isolates from
 Falconiformes and Psittaciformes.";
 RL Arch. Virol. 143:1501-1512(1998).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA John R., Jungmann A., Muller H.;
 RT "Agonoprotein 1a and agnoprotein 1b of avian polyomavirus are apoptotic
 inducers.";
 RL J. Gen. Virol. 0:0-0(2000).
 DR EMBL; AF241169; AAF97886.1; -;
 SQ SEQUENCE 599 AA; 68223 MW; D3D0F04BF6357156 CRC64;

Query Match 2.7%; Score 8; DB 12; Length 599;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 259 RRLTELLG 266
 |||||
 DB 5 RRLTELLG 12

RESULT 34

Q9IBK8 ID Q9IBK8 PRELIMINARY; PRT; 599 AA.

AC Q9IBK8;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE LARGE T ANTIGEN.
 OS Budgerigar fledgling disease virus - 5.
 OC Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Polyomavirus.
 OX NCBI_TaxID=133791;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=88265888; PubMed=2838972;
 RA Rott O., Kroger M., Muller H., Hobom G.;
 RT "The genome of budgerigar fledgling disease virus, an avian
 polyomavirus.";
 RL Virology 165:74-86(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=93155648; PubMed=8381462;
 RA Stoll R., Luo D., Kouwenhoven B., Hobom G., Muller H.;
 RT "Molecular and biological characteristics of avian polyomaviruses:
 isolates from different species of birds indicate that avian
 polyomaviruses form a distinct subgenus within the polyomavirus
 genus.";
 RL J. Gen. Virol. 74:229-237(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=95146945; PubMed=7844527;
 RA Luo D., Muller H., Tang X.B., Hobom G.;
 RT "Early and late pre-mRNA processing of budgerigar fledgling disease
 virus 1: identification of viral RNA 5' and 3' ends and internal
 splice junctions.";
 RL J. Gen. Virol. 76:161-166(1995).
 RN [4]
 RP SEQUENCE FROM N.A.

```

RX MEDLINE=98411646; PubMed=9739329;
RA John R., Muller H.;
RT "Avian polyomavirus in wild birds: genome analysis of isolates from
RT Falconiformes and Psittaciformes.";
RN Arch. Virol. 143:1501-1512(1998).
RN [5]
RP SEQUENCE FROM N.A.
RA John R., Jungmann A., Muller H.;
RT "Agno protein 1a and agno protein 1b of avian polyomavirus are apoptotic
RT inducers.";
RL J. Gen. Virol. 0:0-0(2000).
SQ EMBL: AF241170; AAF97893.1; -.
SQ SEQUENCE 599 AA; 68210 MW; A52104F4A674DFE6 CRC64;

Query Match 2.7%; Score 8; DB 12; Length 599;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 259 RRLTELLG 266
Db 5 RRLTELLG 12

RESULT 35
Q90974 PRELIMINARY; PRT; 644 AA.
AC Q90974: Q91020;
DT 01-JAN-1999 (TrEMBLrel. 09, Created)
DT 01-JAN-1999 (TrEMBLrel. 09, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE MUELLERIAN INHIBITING FACTOR PRECURSOR (MIS) (ANTI-MUELLERIAN HORMONE)
DE (AMH) (MUELLERIAN INHIBITING SUBSTANCE).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WHITE LEGHORN; TISSUE=TESTIS;
RA Eusebe D.C., di Clemente N., Rey R., Piau C., Vigier B., Josso N.,
RA Picard J.-Y.;
RL J. Biol. Chem. 271:4798-4804(1996).
RN [2]
RP REVISIONS.
RC STRAIN-WHITE LEGHORN; TISSUE=TESTIS;
RA Eusebe D.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Neepier M., Lowe R., Galuska S., Hofmann K.J., Smith R.G., Elbrecht A.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS GLYCOPROTEIN, PRODUCED BY THE SERTOLI CELLS OF THE
CC TESTIS, CAUSES REGRESSION OF THE MUELLERIAN DUCT. IT ALSO IS ABLE
CC TO INHIBIT THE GROWTH OF TUMORS DERIVED FROM TISSUES OF MUELLERIAN
CC DUCT ORIGIN.
CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
CC -1- DEVELOPMENTAL STAGE: EXPRESSION PEAKS IN THE TESTIS AT TEN DAYS OF
CC EMBRYONIC LIFE AND DECREASES THEREAFTER WITH ONLY A RELATIVELY
CC LOW AMOUNT PRESENT IN ADULTS. IN FEMALE EMBRYOS, EXPRESSION IS
CC PRESENT IN BOTH OVARIES AT MUCH LOWER LEVELS THAN MALES AND LEVELS
CC ARE EQUAL BETWEEN EIGHT AND TEN DAYS BUT THEREAFTER LEVELS ARE
CC HIGHER IN THE LEFT OVARY AND IN ADULTS, EXPRESSION IS MODERATE IN
CC THE LEFT OVARY.
CC -1- SIMILARITY: TO OTHER GROWTH FACTORS OF THE TGF-BETA FAMILY.
DR EMBL: X89248; CAA61536.1; -.
DR EMBL: U61754; AAB04022.1; -.
DR INTERPRO: IPR001839; -.
DR INTERPRO: IPR002400; -.
DR PFAM: PF000019; TGF-beta; 1.
DR PRINTS: PR00438; GFCYSKNOT.
DR PROSITE: PS00250; TGF_BETA; 1.

```

```

DR PRODUM; PD000357; -. 1.
KW Growth factor; Glycoprotein; Gonadal differentiation; Signal.
FT SIGNAL 1 20
FT PROPEP 21 ?
FT CHAIN ? 644
FT DISULFID 546 610
FT DISULFID 572 641
FT DISULFID 576 643
FT DISULFID 609 609
FT CARBOHYD 49 49
FT CARBOHYD 362 362
FT CARBOHYD 416 416
FT CARBOHYD 537 537
FT VARIANT 65 65
FT VARIANT 194 194
FT CONFLICT 385 385
SQ SEQUENCE 644 AA; 70506 MW; F5D712E2DB93087F CRC64;

Query Match 2.7%; Score 8; DB 13; Length 644;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 200 PLSTRVPG 207
Db 579 PLSTRVPG 586

RESULT 36
Q9P5G4 PRELIMINARY; PRT; 781 AA.
ID Q9P5G4
AC Q9P5G4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE HORMONE-SENSITIVE LIPASE RELATED PROTEIN.
DE B14D6.240.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Sordariales; Sordariaceae; Neurospora.
OC NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Align V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL356173; CAB91733.1; -.
SQ SEQUENCE 781 AA; 85606 MW; BB81C465E1944F5B CRC64;

Query Match 2.7%; Score 8; DB 3; Length 781;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 269 DGALLVRL 276
Db 749 DGALLVRL 756

RESULT 37
Q9L953 PRELIMINARY; PRT; 1013 AA.
ID Q9L953
AC Q9L953;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE SERUM OPACITY FACTOR.
OS SOF63.
OC Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;

```

```
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M63/129 (4/75);
RA Katerov V., Lindgren P.E., Totolian A.A., Schalen C.;
RT "Streptococcal Opacity Factor: A Family of Bifunctional Proteins with
  Lipoproteinase and Fibronectin-Binding Activities.";
RL Curr. Microbiol. 40:149-156(2000).
DR EMBL: AF181974; AAF62872.1; -.
SQ SEQUENCE 1013 AA; 111479 MW; 709B19BA0266F7AE CRC64;

Query Match      2.7%; Score 8; DB 2; Length 1013;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSE 166
DB 109 SASSSSE 116

RESULT 38
O33954 PRELIMINARY; PRT; 4472 AA.
AC O33954;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE TYLACTONE SYNTHASE STARTER MODULE AND MODULES 1 & 2.
GN TYLG.
OS Streptomyces fradiae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1906;
RN [1]
RP SEQUENCE FROM N.A.
RA DeHoff B.S., Sutton K.L., Rostock P.R., Jr.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U78289; AAB66504.1; -.
DR INTERPRO: IPR000255; -.
DR INTERPRO: IPR000794; -.
DR INTERPRO: IPR001227; -.
DR INTERPRO: IPR001993; -.
DR INTERPRO: IPR002106; -.
DR INTERPRO: IPR002198; -.
DR PFAM: PF00106; adh_short; 1.
DR PFAM: PF00109; ketoacyl-synt; 3.
DR PFAM: PF00550; pp-binding; 3.
DR PFAM: PF00698; Acyl_transf; 3.
DR PROSITE: PS00012; PHOSPHOPANTETHEINE; 2.
DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_1.
DR PROSITE: PS00339; AA_TRNA_LIGASE_IL_2; UNKNOWN_3.
DR PROSITE: PS00606; B_KETOACYL_SYNTHASE; 2.
DR PROSITE: PS50075; ACP_DOMAIN; 3.
KW Transferase.
SQ SEQUENCE 4472 AA; 463281 MW; E911D1B9E0C332CF CRC64;

Query Match      2.7%; Score 8; DB 2; Length 4472;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 LRRRLTEL 264
DB 4424 LRRRLTEL 4431

RESULT 39
Q9P7A7 PRELIMINARY; PRT; 42 AA.
AC Q9P7A7;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
```

```
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE HYPOTHETICAL SERINE-RICH PROTEIN (FRAGMENT).
GN SPAC1705.03C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;
OC Schizosaccharomycetaceae; Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972H-;
RA Brown S., Harris D., Wood V., Rajandream M.A., Barrell B.G.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL163529; CAB86946.1; -.
FT NON_TER 42 42
SQ SEQUENCE 42 AA; 3953 MW; 262E24C5F241951C CRC64;

Query Match      2.3%; Score 7; DB 3; Length 42;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSS 165
DB 7 SASSSS 13

RESULT 40
Q9TR13 PRELIMINARY; PRT; 56 AA.
AC Q9TR13;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE SECRETIN (FRAGMENT).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE.
RX MEDLINE=96109189; PubMed=8618828;
RA Bonetto V., Jornvall H., Mutt V., Sillard R.;
RT "Two alternative processing pathways for a preprohormone: a bioactive
  form of secretin.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:11985-11989(1995).
DR HSSP: P01274; IGCN.
DR INTERPRO: IPR000532; -.
DR PFAM: PF00123; hormone2; 1.
DR PROSITE: PS00260; GLUCAGON; 1.
SQ SEQUENCE 56 AA; 6095 MW; B2D9B2DFC3CC5241 CRC64;

Query Match      2.3%; Score 7; DB 6; Length 56;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLIQ 234
DB 47 RLQRLIQ 53

RESULT 41
O76201 PRELIMINARY; PRT; 82 AA.
AC O76201;
DT 01-NOV-1998 (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE NEUROTOXIN TX3-2 PRECURSOR.
GN TX3-2.
OS Phoneutria nigriventer (Brazilian armed spider).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
```

OC Araneomorphae; Entelegynae; Lycosoidea; Ctenidae; Phoneutria.
 OX NCBI_TaxID=6918;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99053416; PubMed=9839681;
 RC TISSUE=VENOM GLAND;
 RA Kalapothakis E., Penaforte C.L., Leao R.M., Cruz J.S., Prado V.F.,
 RA Cordeiro M.N., Diniz C.R., Romano-Silva M.A., Prado M.A.M.,
 RA Gomez M.V., Belrao P.S.L.;
 RT "Cloning, cDNA sequence analysis and patch clamp studies of a toxin
 from the venom of the armed spider (Phoneutria nigriventer).";
 RL Toxicon 36:1971-1980(1998).
 RN [2]
 RP SEQUENCE OF 38-71.
 RC TISSUE=VENOM;
 RX MEDLINE=93190315; PubMed=8446961;
 RA Cordeiro M. do N., De Figueiredo S.G., Valentim A. do C., Diniz C.R.,
 RA von Eickstedt V.R.D., Gilroy J., Richardson M.;
 RT "Purification and amino acid sequences of six Tx3 type neurotoxins
 from the venom of the Brazilian 'armed' spider Phoneutria nigriventer
 (Keys).";
 RL Toxicon 31:35-42(1993).
 CC -1- FUNCTION: ANTAGONIST OF L-TYPE CALCIUM CHANNELS. INDUCES IMMEDIATE
 CC CLOCKWISE GYRATION AND FLACCID PARALYSIS AFTER 6 HOURS AT DOSE
 CC LEVELS OF 5 MICROGRAMS PER MOUSE.
 EMBL; AF015664; AAC26168.1; -;
 KW Venom; Neurotoxin; Signal.
 FT SIGNAL 1 21 POTENTIAL.
 FT PROPEP 22 37
 FT CHAIN 38 71 NEUROTOXIN TX3-2.
 FT PROPEP 72 82
 SQ SEQUENCE 82 AA; 8927 MW; 0CA4E00F5AAAD8B9 CRC64;

Query Match 2.3%; Score 7; DB 5; Length 82;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 EEARAC 105
 DB 33 EEARAC 39
 ID Q28411 PRELIMINARY; PRT; 82 AA.
 AC Q28411;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DE TUMOUR NECROSIS FACTOR RECEPTOR P80 (FRAGMENT).
 OS Felis silvestris catus (Cat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
 OX NCBI_TaxID=9685;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Duthie S., Nasir L., Eckersall P.D.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U51429; AAB01940.1; -;
 FT NON_TER 1 1
 FT NON_TER 82 82
 SQ SEQUENCE 82 AA; 7865 MW; E2F9DA7301EAD1E CRC64;

Query Match 2.3%; Score 7; DB 6; Length 82;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 ASSSSSE 166
 DB 31 ASSSSSE 37

RESULT 43
 Q9S9T1
 ID Q9S9T1 PRELIMINARY; PRT; 87 AA.
 AC Q9S9T1;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE T32N4-14 PROTEIN (AT4G05070 PROTEIN).
 GN T32N4-14 OR AT4G05070.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLOMBIA;
 RA Layman D., Du H., Elliot G., Shafer S.;
 RT "The sequence of A. thaliana T32N4.";
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLOMBIA;
 RA Washu;
 RT "The A. thaliana Genome Sequencing Project.";
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLOMBIA;
 RA Waterston R.;
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vil D.M.,
 RA Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,
 RA Shekher M., Schütz K., See U.H., Swaby I., Habermann K., Dedhia N.N.,
 RA Newes H.W., Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF162444; AAD48981.1; -;
 DR EMBL; AL161502; CAB81049.1; -;
 SQ SEQUENCE 87 AA; 9307 MW; 20C4BEF8EEB6025F CRC64;

Query Match 2.3%; Score 7; DB 10; Length 87;
 Best Local Similarity 100.0%; Pred. No. 56;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
 DB 41 SASSSSS 47

RESULT 44
 Q9VQP7
 ID Q9VQP7 PRELIMINARY; PRT; 90 AA.
 AC Q9VQP7;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE CG15411 PROTEIN.
 GN CG15411.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.


```
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Brill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Moberly C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveril J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003580; AAF51118.1; -.
DR FLYBASE; FBgn0040712; CG15411.
SQ SEQUENCE 90 AA; 9256 MW; 0713C45E3A2D9858 CRC64;

Query Match 2.3%; Score 7; DB 5; Length 90;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 159 SASSSSS 165
Db 17 SASSSSS 23

RESULT 45
Q9NFHO PRELIMINARY; PRT; 98 AA.
AC Q9NFH0;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE HYPOTHETICAL 9.6 KDA PROTEIN (FRAGMENT).
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RA Spielmann T., Beck H.;
RT "Analysis of stage specific transcription in Plasmodium falciparum
RT reveals a set of genes exclusively transcribed in ring stage
RT parasites.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
```

RN SEQUENCE FROM N.A.
 RP STRAIN=KS3-2;
 RX MEDLINE=92145774; PubMed=1723663;
 RT Winkler M., Kueck U.;
 RT "The group IIB intron from the green alga *Scenedesmus obliquus*
 RT mitochondrion: molecular characterization of the in vitro splicing
 RT products.";
 RL Curr. Genet. 20:495-502(1991).
 RP SEQUENCE FROM N.A.
 RP STRAIN=KS3-2;
 RA Kueck U., Jekosch K., Holzamer P.;
 RT "DNA sequence analysis of the complete mitochondrial genome of the
 RT green alga *Scenedesmus obliquus*: Evidence for UAG being a leucine and
 RT UCA being a non-sense codon.";
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RP SEQUENCE FROM N.A.
 RP STRAIN=UTEX 78;
 RA Nedelcu A.M., Lee R.W., Lemieux C., Gray M.W., Burger G.;
 RT "The complete mitochondrial DNA sequence of *Scenedesmus obliquus*
 RT reflects an intermediate stage in the evolution of the green algal
 RT mitochondrial genome.";
 RL Genome Res. 10:819-831(2000).
 RP SEQUENCE FROM N.A.
 RP STRAIN=UTEX 78;
 RA Burger G.; (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; X17375; CAB90362.1; -.
 DR EMBL; AF204057; AAF72054.1; -.
 KW Mitochondrion; Oxidoreductase.
 SQ SEQUENCE 100 AA; 10841 MW; 6F3C3C9533D1284A CRC64;

Query Match 2.3%; Score 7; DB 8; Length 100;
 Best Local Similarity 100.0%; Pred. No. 63;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 175 TALGLAL 181
 Db |||||
 73 TALGLAL 79

RESULT 48
 ID Q9PEH3 PRELIMINARY; PRT; 106 AA.
 AC Q9PEH3;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE HYPOTHETICAL PROTEIN XF1055.
 GN XF1055.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 OC Xylella.
 OX NCBI_TaxID=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9A5C;
 RX MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,

RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silvestri M.L.Z., Silveira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchioka M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen *Xylella fastidiosa*.";
 RL Nature 406:151-157(2000).
 DR EMBL; AE003942; AAF83865.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 106 AA; 12175 MW; 9E048C950D8A6959 CRC64;

Query Match 2.3%; Score 7; DB 2; Length 106;
 Best Local Similarity 100.0%; Pred. No. 66;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 RDAETGE 45
 Db |||||
 87 RDAETGE 93

RESULT 49
 ID Q9YAE2 PRELIMINARY; PRT; 112 AA.
 AC Q9YAE2;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DE 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE HYPOTHETICAL 12.0 KDA PROTEIN APE1997.
 GN APE1997.
 OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;
 OC Aeropyrum.
 OX NCBI_TaxID=56636;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K1;
 RX MEDLINE=99310339; PubMed=10382966;
 RA Kawarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
 RA Jin-no K., Takahashi M., Sekine M., Baba S., Ankai A., Kosugi H.,
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
 RA Yamazaki J., Kushida N., Oguchi A., Aoki K., Kubota K., Nakamura Y.,
 RA Nomura N., Sako Y., Kikuchi H.;
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 RT crenarchaeon, *Aeropyrum pernix* K1.";
 RL DNA Res. 6:83-101(1999).
 DR EMBL; AP000063; BAA81007.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 112 AA; 11993 MW; 3554696AD1928C3D CRC64;

Query Match 2.3%; Score 7; DB 1; Length 112;
 Best Local Similarity 100.0%; Pred. No. 69;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
 Db |||||
 3 SASSSSS 9

RESULT 50
 Q29173

ID Q29173 PRELIMINARY; PRT; 115 AA.
AC Q29173;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE CYTOPLASMIC BETA-ACTIN (FRAGMENT).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SMALL INTESTINE;
RX MEDLINE=96327607; PubMed=8672129;
RA Winteroe A.K., Fredholm M., Davies W.;
RT "Evaluation and characterization of a porcine small intestine cDNA
library: analysis of 839 clones."
RL Mamm. Genome 7:509-517(1996).
DR EMBL; F14786; CAA23255.1; -;
DR INTERPRO; IPR00279; -;
DR PFAM; PF00022; actin; 1.
FT NON_TER 1
FT NON_TER 115
SQ SEQUENCE 115 AA; 13566 MW; 1891FC8A9FBAL49B CRC64;

Query Match 2.3%; Score 7; DB 6; Length 115;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 ALPALLP 23
Db 56 ALPALLP 62

Search completed: January 30, 2001, 17:10:38
Job time: 257 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 16:50:50 ; Search time 22.54 Seconds
(without alignments)
239.002 Million cell updates/sec

Title: US-09-518-931-2
Perfect score: 300
Sequence: 1 MRALEGPGLSLCLLVLALPA.....RVARMPLGERSVREPLVH 300

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 174772 seqs, 17957048 residues

Word size : 6

Total number of hits satisfying chosen parameters: 663

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6_COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/PCITUS_COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	300	100.0	300	2	US-08-794-796-2
2	8	2.7	28	3	US-08-938-548B-4
3	8	2.7	28	3	US-08-938-548B-9
4	8	2.7	28	3	US-08-938-548B-12
5	8	2.7	123	3	US-08-938-548B-10
6	8	2.7	130	3	US-08-938-548B-6
7	8	2.7	131	3	US-08-938-548B-2
8	8	2.7	1172	1	US-08-313-288B-19
9	8	2.7	4472	2	US-08-804-227C-2
10	7	2.3	26	1	US-07-776-272-25
11	7	2.3	27	1	US-07-924-054-10
12	7	2.3	27	1	US-08-062-472B-43
13	7	2.3	27	1	US-08-519-180-6
14	7	2.3	27	2	US-08-818-253-36
15	7	2.3	69	2	US-08-583-569-1
16	7	2.3	150	1	US-08-374-843B-6
17	7	2.3	150	1	US-08-374-843B-10
18	7	2.3	150	2	US-08-905-420-6
19	7	2.3	150	2	US-08-905-420-10
20	7	2.3	181	1	US-08-185-432-6
21	7	2.3	235	3	US-09-066-408-12
22	7	2.3	283	4	PCT-US96-12374-2
23	7	2.3	315	3	US-08-965-903B-8
24	7	2.3	393	3	US-08-888-429A-21
25	7	2.3	410	1	US-08-123-343A-5
26	7	2.3	410	1	US-08-123-343A-7
27	7	2.3	430	1	US-08-318-947A-16
28	7	2.3	430	2	US-08-795-303-16
29	7	2.3	430	2	US-08-795-303-16
30	7	2.3	430	2	US-08-795-303-16
31	7	2.3	430	2	US-08-795-303-16
32	7	2.3	430	2	US-08-795-303-16
33	7	2.3	430	2	US-08-795-303-16
34	7	2.3	430	2	US-08-795-303-16
35	7	2.3	430	2	US-08-795-303-16
36	7	2.3	430	2	US-08-795-303-16
37	7	2.3	430	2	US-08-795-303-16
38	7	2.3	430	2	US-08-795-303-16
39	7	2.3	430	2	US-08-795-303-16
40	7	2.3	430	2	US-08-795-303-16
41	7	2.3	430	2	US-08-795-303-16
42	7	2.3	430	2	US-08-795-303-16
43	7	2.3	430	2	US-08-795-303-16
44	7	2.3	430	2	US-08-795-303-16
45	7	2.3	430	2	US-08-795-303-16
46	7	2.3	430	2	US-08-795-303-16
47	7	2.3	430	2	US-08-795-303-16
48	7	2.3	430	2	US-08-795-303-16
49	6	2.0	9	2	US-08-482-651-10
50	6	2.0	24	2	US-08-752-891-7
51	6	2.0	24	2	US-09-144-178-7
52	6	2.0	24	3	US-09-406-854-7
53	6	2.0	31	1	US-08-418-716A-3
54	6	2.0	36	3	US-08-984-151-4
55	6	2.0	38	1	US-08-176-500-56
56	6	2.0	38	1	US-08-471-052A-56
57	6	2.0	38	1	US-08-189-331-56
58	6	2.0	38	2	US-08-471-939-56
59	6	2.0	38	2	US-08-471-800-56
60	6	2.0	38	2	US-08-471-068-56
61	6	2.0	43	2	US-08-169-948B-20
62	6	2.0	43	2	US-08-448-873-20
63	6	2.0	44	3	US-08-888-429A-33
64	6	2.0	45	1	US-08-278-089A-25
65	6	2.0	45	2	US-08-388-957A-24
66	6	2.0	54	1	US-08-605-053-1
67	6	2.0	60	2	US-08-775-009-39
68	6	2.0	66	2	US-08-588-258B-7
69	6	2.0	66	3	US-08-460-505-7
70	6	2.0	66	4	PCT-US96-08295-7
71	6	2.0	70	3	US-09-188-930-131
72	6	2.0	84	3	US-09-107-858-27
73	6	2.0	100	1	US-08-473-981A-11
74	6	2.0	100	2	US-08-474-087-11
75	6	2.0	103	2	US-08-273-146-71
76	6	2.0	126	3	US-09-053-197A-11
77	6	2.0	126	3	US-08-513-974B-28
78	6	2.0	127	1	US-08-266-451B-28
79	6	2.0	127	2	US-08-748-725-28
80	6	2.0	130	2	US-08-630-822A-90
81	6	2.0	130	2	US-09-005-069-90
82	6	2.0	139	3	US-09-053-197A-40
83	6	2.0	142	2	US-08-164-292B-20
84	6	2.0	142	3	US-08-845-623-20
85	6	2.0	142	3	US-08-815-927-20
86	6	2.0	144	3	US-08-513-974B-366
87	6	2.0	144	3	US-08-513-974B-369
88	6	2.0	148	4	PCT-US95-07135-2
89	6	2.0	149	2	US-08-461-030C-2
90	6	2.0	149	3	US-08-744-138-2
91	6	2.0	150	2	US-08-851-188-1
92	6	2.0	150	2	US-08-851-188-3
93	6	2.0	159	2	US-08-851-188-4
94	6	2.0	162	1	US-08-266-451B-29
95	6	2.0	162	2	US-08-748-725-29
96	6	2.0	169	3	US-08-833-167-58
97	6	2.0	171	3	US-08-833-167-43
98	6	2.0	171	3	US-08-833-167-45
99	6	2.0	171	3	US-08-833-167-44
100	6	2.0	171	3	US-08-833-167-47
101	6	2.0	171	3	US-08-833-167-57

102	6	2.0	171	3	US-08-833-167-59	Sequence 59, Appl	175	1	US-08-010-099-104	Sequence 104, App
103	6	2.0	171	3	US-08-833-167-60	Sequence 60, Appl	176	1	US-08-010-099-105	Sequence 105, App
104	6	2.0	173	1	US-08-354-456A-7	Sequence 7, Appl	177	1	US-08-010-099-106	Sequence 106, App
105	6	2.0	174	1	US-07-734-225A-43	Sequence 43, Appl	178	1	US-08-010-099-107	Sequence 107, App
106	6	2.0	174	1	US-07-692-935B-46	Sequence 46, Appl	179	1	US-08-010-099-108	Sequence 108, App
107	6	2.0	174	1	US-08-010-099-82	Sequence 82, Appl	180	1	US-08-010-099-109	Sequence 109, App
108	6	2.0	174	1	US-08-225-224-5	Sequence 5, Appl	181	1	US-08-010-099-110	Sequence 110, App
109	6	2.0	174	1	US-08-434-411-2	Sequence 2, Appl	182	1	US-08-167-721-1	Sequence 1, Appl
110	6	2.0	174	1	US-08-434-402-2	Sequence 43, Appl	183	1	US-08-428-732-6	Sequence 6, Appl
111	6	2.0	174	1	US-08-488-457-43	Sequence 43, Appl	184	1	US-08-448-716-2	Sequence 2, Appl
112	6	2.0	174	1	US-08-448-716-82	Sequence 82, Appl	185	1	US-08-448-716-67	Sequence 67, Appl
113	6	2.0	174	1	US-08-783-288-2	Sequence 2, Appl	186	1	US-08-448-716-68	Sequence 68, Appl
114	6	2.0	174	2	US-08-431-459A-31	Sequence 31, Appl	187	1	US-08-448-716-69	Sequence 69, Appl
115	6	2.0	174	2	US-08-890-640-2	Sequence 2, Appl	188	1	US-08-448-716-70	Sequence 70, Appl
116	6	2.0	174	3	US-08-722-258-5	Sequence 5, Appl	189	1	US-08-448-716-71	Sequence 71, Appl
117	6	2.0	174	3	US-08-833-167-1	Sequence 1, Appl	190	1	US-08-448-716-72	Sequence 72, Appl
118	6	2.0	174	3	US-08-833-167-48	Sequence 48, Appl	191	1	US-08-448-716-73	Sequence 73, Appl
119	6	2.0	174	3	US-08-833-167-49	Sequence 49, Appl	192	1	US-08-448-716-74	Sequence 74, Appl
120	6	2.0	174	3	US-08-833-167-50	Sequence 50, Appl	193	1	US-08-448-716-75	Sequence 75, Appl
121	6	2.0	174	3	US-08-833-167-51	Sequence 51, Appl	194	1	US-08-448-716-76	Sequence 76, Appl
122	6	2.0	174	3	US-08-833-167-52	Sequence 52, Appl	195	1	US-08-448-716-77	Sequence 77, Appl
123	6	2.0	174	3	US-08-833-167-53	Sequence 53, Appl	196	1	US-08-448-716-78	Sequence 78, Appl
124	6	2.0	174	3	US-08-833-167-54	Sequence 54, Appl	197	1	US-08-448-716-79	Sequence 79, Appl
125	6	2.0	174	3	US-08-833-167-55	Sequence 55, Appl	198	1	US-08-448-716-80	Sequence 80, Appl
126	6	2.0	174	3	US-08-833-167-56	Sequence 56, Appl	199	1	US-08-448-716-81	Sequence 81, Appl
127	6	2.0	174	3	US-08-833-167-95	Sequence 95, Appl	200	1	US-08-448-716-83	Sequence 83, Appl
128	6	2.0	174	3	US-08-833-167-96	Sequence 96, Appl	201	1	US-08-448-716-84	Sequence 84, Appl
129	6	2.0	174	3	US-08-833-167-97	Sequence 97, Appl	202	1	US-08-448-716-85	Sequence 85, Appl
130	6	2.0	174	3	US-08-833-167-98	Sequence 98, Appl	203	1	US-08-448-716-86	Sequence 86, Appl
131	6	2.0	174	3	US-08-833-167-100	Sequence 100, App	204	1	US-08-448-716-87	Sequence 87, Appl
132	6	2.0	174	3	US-08-833-167-101	Sequence 101, App	205	1	US-08-448-716-88	Sequence 88, Appl
133	6	2.0	174	3	US-08-833-167-102	Sequence 102, App	206	1	US-08-448-716-89	Sequence 89, Appl
134	6	2.0	174	3	US-08-833-167-103	Sequence 103, App	207	1	US-08-448-716-90	Sequence 90, Appl
135	6	2.0	174	4	PCT-US95-04468-5	Sequence 5, Appl	208	1	US-08-448-716-91	Sequence 91, Appl
136	6	2.0	174	5	5194592-26	Patent No. 5194592	209	1	US-08-448-716-92	Sequence 92, Appl
137	6	2.0	174	5	5194592-27	Patent No. 5194592	210	1	US-08-448-716-93	Sequence 93, Appl
138	6	2.0	175	1	US-08-010-099-2	Sequence 2, Appl	211	1	US-08-448-716-94	Sequence 94, Appl
139	6	2.0	175	1	US-08-010-099-67	Sequence 67, Appl	212	1	US-08-448-716-95	Sequence 95, Appl
140	6	2.0	175	1	US-08-010-099-68	Sequence 68, Appl	213	1	US-08-448-716-96	Sequence 96, Appl
141	6	2.0	175	1	US-08-010-099-69	Sequence 69, Appl	214	1	US-08-448-716-97	Sequence 97, Appl
142	6	2.0	175	1	US-08-010-099-70	Sequence 70, Appl	215	1	US-08-448-716-98	Sequence 98, Appl
143	6	2.0	175	1	US-08-010-099-71	Sequence 71, Appl	216	1	US-08-448-716-99	Sequence 99, Appl
144	6	2.0	175	1	US-08-010-099-72	Sequence 72, Appl	217	1	US-08-448-716-100	Sequence 100, App
145	6	2.0	175	1	US-08-010-099-73	Sequence 73, Appl	218	1	US-08-448-716-101	Sequence 101, App
146	6	2.0	175	1	US-08-010-099-74	Sequence 74, Appl	219	1	US-08-448-716-102	Sequence 102, App
147	6	2.0	175	1	US-08-010-099-75	Sequence 75, Appl	220	1	US-08-448-716-103	Sequence 103, App
148	6	2.0	175	1	US-08-010-099-76	Sequence 76, Appl	221	1	US-08-448-716-104	Sequence 104, App
149	6	2.0	175	1	US-08-010-099-77	Sequence 77, Appl	222	1	US-08-448-716-105	Sequence 105, App
150	6	2.0	175	1	US-08-010-099-78	Sequence 78, Appl	223	1	US-08-448-716-106	Sequence 106, App
151	6	2.0	175	1	US-08-010-099-79	Sequence 79, Appl	224	1	US-08-448-716-107	Sequence 107, App
152	6	2.0	175	1	US-08-010-099-80	Sequence 80, Appl	225	1	US-08-448-716-108	Sequence 108, App
153	6	2.0	175	1	US-08-010-099-81	Sequence 81, Appl	226	1	US-08-448-716-109	Sequence 109, App
154	6	2.0	175	1	US-08-010-099-83	Sequence 83, Appl	227	1	US-08-448-716-110	Sequence 110, App
155	6	2.0	175	1	US-08-010-099-84	Sequence 84, Appl	228	1	US-08-321-510-2	Sequence 2, Appl
156	6	2.0	175	1	US-08-010-099-85	Sequence 85, Appl	229	1	US-08-570-943-1	Sequence 1, Appl
157	6	2.0	175	1	US-08-010-099-86	Sequence 86, Appl	230	1	US-08-879-760-2	Sequence 2, Appl
158	6	2.0	175	1	US-08-010-099-87	Sequence 87, Appl	231	1	US-08-865-297-6	Sequence 6, Appl
159	6	2.0	175	1	US-08-010-099-88	Sequence 88, Appl	232	1	PCT-US95-01729-2	Sequence 2, Appl
160	6	2.0	175	1	US-08-010-099-89	Sequence 89, Appl	233	1	US-08-469-318-161	Sequence 161, App
161	6	2.0	175	1	US-08-010-099-90	Sequence 90, Appl	234	1	US-08-469-318-162	Sequence 162, App
162	6	2.0	175	1	US-08-010-099-91	Sequence 91, Appl	235	1	US-08-468-609A-161	Sequence 161, App
163	6	2.0	175	1	US-08-010-099-92	Sequence 92, Appl	236	1	US-08-468-609A-162	Sequence 162, App
164	6	2.0	175	1	US-08-010-099-93	Sequence 93, Appl	237	1	PCT-US94-12873-26	Sequence 26, Appl
165	6	2.0	175	1	US-08-010-099-94	Sequence 94, Appl	238	1	PCT-US95-01185-161	Sequence 161, App
166	6	2.0	175	1	US-08-010-099-95	Sequence 95, Appl	239	1	PCT-US95-01185-162	Sequence 162, App
167	6	2.0	175	1	US-08-010-099-96	Sequence 96, Appl	240	1	US-08-338-793D-41	Sequence 41, Appl
168	6	2.0	175	1	US-08-010-099-97	Sequence 97, Appl	241	1	US-08-338-793D-55	Sequence 55, Appl
169	6	2.0	175	1	US-08-010-099-98	Sequence 98, Appl	242	1	US-08-431-459A-30	Sequence 30, Appl
170	6	2.0	175	1	US-08-010-099-99	Sequence 99, Appl	243	1	US-08-797-689-14	Sequence 14, Appl
171	6	2.0	175	1	US-08-010-099-100	Sequence 100, App	244	1	US-09-122-443-8	Sequence 8, Appl
172	6	2.0	175	1	US-08-010-099-101	Sequence 101, App	245	1	US-08-833-167-116	Sequence 116, App
173	6	2.0	175	1	US-08-010-099-102	Sequence 102, App	246	1	US-08-833-167-117	Sequence 117, App
174	6	2.0	175	1	US-08-010-099-103	Sequence 103, App	247	1	US-08-833-167-118	Sequence 118, App

248	6	2.0	177	3	US-08-833-167-119	Sequence 119, App	321	6	2.0	307	3	US-08-469-318-146	Sequence 146, App
249	6	2.0	177	3	US-08-833-167-120	Sequence 120, App	322	6	2.0	307	3	US-08-469-318-147	Sequence 147, App
250	6	2.0	177	3	US-08-833-167-121	Sequence 121, App	323	6	2.0	307	3	US-08-469-318-152	Sequence 152, App
251	6	2.0	177	3	US-08-833-167-122	Sequence 122, App	324	6	2.0	307	3	US-08-469-318-158	Sequence 158, App
252	6	2.0	177	3	US-08-833-167-123	Sequence 123, App	325	6	2.0	307	3	US-08-469-318-159	Sequence 159, App
253	6	2.0	177	3	US-08-833-167-124	Sequence 124, App	326	6	2.0	307	3	US-08-468-609A-121	Sequence 121, App
254	6	2.0	177	3	US-08-833-167-125	Sequence 125, App	327	6	2.0	307	3	US-08-468-609A-122	Sequence 122, App
255	6	2.0	177	3	US-08-833-167-126	Sequence 126, App	328	6	2.0	307	3	US-08-468-609A-123	Sequence 123, App
256	6	2.0	177	3	US-08-833-167-127	Sequence 127, App	329	6	2.0	307	3	US-08-468-609A-124	Sequence 124, App
257	6	2.0	177	3	US-08-833-167-128	Sequence 128, App	330	6	2.0	307	3	US-08-468-609A-134	Sequence 134, App
258	6	2.0	183	5	5168049-4	Patent No. 5168049	331	6	2.0	307	3	US-08-468-609A-135	Sequence 135, App
259	6	2.0	184	4	PCT-US94-12873-23	Sequence 23, Appl	332	6	2.0	307	3	US-08-468-609A-146	Sequence 146, App
260	6	2.0	203	3	US-09-106-182-3	Sequence 3, Appl	333	6	2.0	307	3	US-08-468-609A-147	Sequence 147, App
261	6	2.0	204	1	US-08-792-019B-10	Sequence 10, Appl	334	6	2.0	307	3	US-08-468-609A-152	Sequence 152, App
262	6	2.0	204	3	US-08-988-819-10	Sequence 10, Appl	335	6	2.0	307	3	US-08-468-609A-158	Sequence 158, App
263	6	2.0	204	3	US-09-016-534-10	Sequence 10, Appl	336	6	2.0	307	3	US-08-468-609A-159	Sequence 159, App
264	6	2.0	205	2	US-08-775-009-37	Sequence 37, Appl	337	6	2.0	307	4	PCT-US95-01185-121	Sequence 121, App
265	6	2.0	206	3	US-08-513-974B-27	Sequence 27, Appl	338	6	2.0	307	4	PCT-US95-01185-122	Sequence 122, App
266	6	2.0	211	2	US-08-708-958-2	Sequence 1, Appl	339	6	2.0	307	4	PCT-US95-01185-123	Sequence 123, App
267	6	2.0	211	2	US-08-708-958-2	Sequence 2, Appl	340	6	2.0	307	4	PCT-US95-01185-124	Sequence 124, App
268	6	2.0	215	1	US-08-266-451B-27	Sequence 27, Appl	341	6	2.0	307	4	PCT-US95-01185-134	Sequence 134, App
269	6	2.0	215	2	US-08-748-725-27	Sequence 27, Appl	342	6	2.0	307	4	PCT-US95-01185-135	Sequence 135, App
270	6	2.0	218	1	US-08-463-115-92	Sequence 92, Appl	343	6	2.0	307	4	PCT-US95-01185-146	Sequence 146, App
271	6	2.0	218	3	US-08-465-388-92	Sequence 92, Appl	344	6	2.0	307	4	PCT-US95-01185-147	Sequence 147, App
272	6	2.0	218	3	US-08-875-233-2	Sequence 2, Appl	345	6	2.0	307	4	PCT-US95-01185-152	Sequence 152, App
273	6	2.0	218	3	US-08-875-233-6	Sequence 6, Appl	346	6	2.0	307	4	PCT-US95-01185-158	Sequence 158, App
274	6	2.0	219	1	US-08-463-115-91	Sequence 91, Appl	347	6	2.0	307	4	PCT-US95-01185-159	Sequence 159, App
275	6	2.0	219	1	US-08-465-388-91	Sequence 91, Appl	348	6	2.0	310	3	US-08-651-136C-22	Sequence 22, Appl
276	6	2.0	220	2	US-08-840-683-8	Sequence 8, Appl	349	6	2.0	317	2	US-08-555-722-8	Sequence 8, Appl
277	6	2.0	223	3	US-08-513-974B-315	Sequence 315, App	350	6	2.0	317	2	US-08-605-284B-17	Sequence 17, Appl
278	6	2.0	223	3	US-08-513-974B-364	Sequence 364, App	351	6	2.0	320	3	US-08-530-165-7	Sequence 7, Appl
279	6	2.0	223	3	US-08-513-974B-368	Sequence 368, App	352	6	2.0	322	3	US-08-469-318-128	Sequence 128, App
280	6	2.0	226	5	5498600-2	Patent No. 5498600	353	6	2.0	322	3	US-08-469-318-129	Sequence 129, App
281	6	2.0	235	1	US-08-015-985-5	Sequence 5, Appl	354	6	2.0	322	3	US-08-469-318-130	Sequence 130, App
282	6	2.0	237	2	US-08-469-537A-85	Sequence 85, Appl	355	6	2.0	322	3	US-08-469-318-138	Sequence 138, App
283	6	2.0	241	5	5175255-2	Patent No. 5175255	356	6	2.0	322	3	US-08-469-318-149	Sequence 149, App
284	6	2.0	241	5	5175255-8	Patent No. 5175255	357	6	2.0	322	3	US-08-469-318-150	Sequence 150, App
285	6	2.0	245	1	US-07-876-284-2	Sequence 2, Appl	358	6	2.0	322	3	US-08-469-318-154	Sequence 154, App
286	6	2.0	245	1	US-08-276-151-9	Sequence 9, Appl	359	6	2.0	322	3	US-08-469-318-156	Sequence 156, App
287	6	2.0	246	1	US-07-887-072B-4	Sequence 4, Appl	360	6	2.0	322	3	US-08-469-318-157	Sequence 157, App
288	6	2.0	246	1	US-08-276-151-7	Sequence 7, Appl	361	6	2.0	322	3	US-08-468-609A-128	Sequence 128, App
289	6	2.0	246	1	US-08-466-444-4	Sequence 4, Appl	362	6	2.0	322	3	US-08-468-609A-138	Sequence 138, App
290	6	2.0	248	1	US-08-266-451B-2	Sequence 2, Appl	363	6	2.0	322	3	US-08-468-609A-130	Sequence 130, App
291	6	2.0	248	1	US-08-748-725-2	Sequence 2, Appl	364	6	2.0	322	3	US-08-468-609A-138	Sequence 138, App
292	6	2.0	249	1	US-07-887-072B-2	Sequence 2, Appl	365	6	2.0	322	3	US-08-468-609A-149	Sequence 149, App
293	6	2.0	249	1	US-08-466-444-2	Sequence 2, Appl	366	6	2.0	322	3	US-08-468-609A-150	Sequence 150, App
294	6	2.0	250	1	US-08-133-804-2	Sequence 2, Appl	367	6	2.0	322	3	US-08-468-609A-154	Sequence 154, App
295	6	2.0	250	1	US-08-461-184-8	Sequence 8, Appl	368	6	2.0	322	3	US-08-468-609A-156	Sequence 156, App
296	6	2.0	250	1	US-08-463-675-8	Sequence 8, Appl	369	6	2.0	322	3	US-08-468-609A-157	Sequence 157, App
297	6	2.0	250	1	US-08-464-589-8	Sequence 8, Appl	370	6	2.0	322	4	PCT-US95-01185-128	Sequence 128, App
298	6	2.0	250	1	US-08-461-838-2	Sequence 2, Appl	371	6	2.0	322	4	PCT-US95-01185-129	Sequence 129, App
299	6	2.0	250	2	US-08-461-386-2	Sequence 2, Appl	372	6	2.0	322	4	PCT-US95-01185-130	Sequence 130, App
300	6	2.0	253	2	US-08-685-992-11	Sequence 11, Appl	373	6	2.0	322	4	PCT-US95-01185-138	Sequence 138, App
301	6	2.0	253	2	US-09-144-925-11	Sequence 11, Appl	374	6	2.0	322	4	PCT-US95-01185-149	Sequence 149, App
302	6	2.0	253	4	PCT-US96-01314-53	Sequence 53, Appl	375	6	2.0	322	4	PCT-US95-01185-150	Sequence 150, App
303	6	2.0	269	3	US-08-759-463-2	Sequence 2, Appl	376	6	2.0	322	4	PCT-US95-01185-154	Sequence 154, App
304	6	2.0	275	3	US-08-036-987A-7	Sequence 7, Appl	377	6	2.0	322	4	PCT-US95-01185-156	Sequence 156, App
305	6	2.0	281	1	US-08-040-548-3	Sequence 3, Appl	378	6	2.0	322	4	PCT-US95-01185-157	Sequence 157, App
306	6	2.0	281	1	US-08-466-344-3	Sequence 3, Appl	379	6	2.0	330	1	US-08-410-167A-4	Sequence 4, Appl
307	6	2.0	282	1	US-08-445-847A-1	Sequence 1, Appl	380	6	2.0	330	2	US-08-898-560-1	Sequence 1, Appl
308	6	2.0	287	1	US-07-971-092-2	Sequence 2, Appl	381	6	2.0	335	3	US-08-469-318-143	Sequence 143, App
309	6	2.0	287	5	5198342-2	Patent No. 5198342	382	6	2.0	335	3	PCT-US95-01185-143	Sequence 143, App
310	6	2.0	292	1	US-08-036-210-12	Sequence 12, Appl	383	6	2.0	335	4	PCT-US95-01185-143	Sequence 143, App
311	6	2.0	292	2	US-08-449-609-12	Sequence 12, Appl	384	6	2.0	337	3	US-08-469-318-148	Sequence 148, App
312	6	2.0	297	3	US-09-187-049-10	Sequence 10, Appl	385	6	2.0	337	3	US-08-469-318-148	Sequence 148, App
313	6	2.0	299	2	US-08-773-608A-2	Sequence 2, Appl	386	6	2.0	337	3	PCT-US95-01185-148	Sequence 148, App
314	6	2.0	303	2	US-08-961-539-2	Sequence 2, Appl	387	6	2.0	342	1	PCT-US95-01185-148	Sequence 9, Appl
315	6	2.0	307	3	US-08-469-318-121	Sequence 121, App	388	6	2.0	342	4	PCT-US93-08528-9	Sequence 9, Appl
316	6	2.0	307	3	US-08-469-318-122	Sequence 122, App	389	6	2.0	346	1	US-08-119-773-5	Sequence 5, Appl
317	6	2.0	307	3	US-08-469-318-123	Sequence 123, App	390	6	2.0	346	2	US-08-602-359A-34	Sequence 34, Appl
318	6	2.0	307	3	US-08-469-318-124	Sequence 124, App	391	6	2.0	348	3	US-09-248-528-7	Sequence 7, Appl
319	6	2.0	307	3	US-08-469-318-134	Sequence 134, App	392	6	2.0	349	2	US-08-494-907-12	Sequence 12, Appl
320	6	2.0	307	3	US-08-469-318-135	Sequence 135, App	393	6	2.0	349	3	US-08-469-318-139	Sequence 139, App

394	6	2.0	349	3	US-08-469-318-151	Sequence 151, App	467	6	2.0	531	2	US-08-752-633-3	Sequence 3, Appl
395	6	2.0	349	3	US-08-468-609A-139	Sequence 139, App	468	6	2.0	531	4	PCT-US95-04886-3	Sequence 3, Appl
396	6	2.0	349	3	US-08-468-609A-151	Sequence 151, App	469	6	2.0	532	1	US-07-618-286-1	Sequence 1, Appl
397	6	2.0	349	4	PCT-US95-01185-139	Sequence 139, App	470	6	2.0	532	2	US-08-196-003-3	Sequence 3, Appl
398	6	2.0	349	4	PCT-US95-01185-151	Sequence 151, App	471	6	2.0	532	1	US-08-483-389-118	Sequence 118, App
399	6	2.0	349	4	PCT-US96-10986-12	Sequence 12, Appl	472	6	2.0	532	2	US-08-689-870-12	Sequence 12, Appl
400	6	2.0	357	1	US-08-119-773-2	Sequence 2, Appl	473	6	2.0	532	3	US-08-933-824-3	Sequence 3, Appl
401	6	2.0	357	1	US-08-119-773-4	Sequence 4, Appl	474	6	2.0	532	5	5284931-2	Patent No. 5284931
402	6	2.0	357	1	US-08-119-773-6	Sequence 6, Appl	475	6	2.0	533	1	US-08-040-548-1	Sequence 1, Appl
403	6	2.0	357	1	US-08-119-773-36	Sequence 36, Appl	476	6	2.0	533	1	US-08-466-344-1	Sequence 1, Appl
404	6	2.0	360	3	US-08-839-437-7	Sequence 7, Appl	477	6	2.0	533	5	5206152-2	Patent No. 5206152
405	6	2.0	370	2	US-08-837-593-7	Sequence 7, Appl	478	6	2.0	534	1	US-08-106-761-2	Sequence 2, Appl
406	6	2.0	370	3	US-08-513-974B-26	Sequence 26, Appl	479	6	2.0	565	2	US-08-892-770-6	Sequence 6, Appl
407	6	2.0	370	3	US-08-513-974B-323	Sequence 323, App	480	6	2.0	566	1	US-08-335-865J-9	Sequence 9, Appl
408	6	2.0	371	2	US-08-837-593-6	Sequence 6, Appl	481	6	2.0	568	1	US-08-320-559-30	Sequence 30, Appl
409	6	2.0	375	2	US-08-837-593-5	Sequence 5, Appl	482	6	2.0	568	3	US-08-545-860D-30	Sequence 30, Appl
410	6	2.0	389	3	US-08-888-429A-20	Sequence 20, Appl	483	6	2.0	568	4	PCT-US94-0496-30	Sequence 30, Appl
411	6	2.0	393	2	US-08-990-379-4	Sequence 4, Appl	484	6	2.0	571	2	US-08-892-770-5	Sequence 5, Appl
412	6	2.0	398	1	US-08-091-519-2	Sequence 2, Appl	485	6	2.0	587	1	US-07-955-905A-23	Sequence 23, Appl
413	6	2.0	398	1	US-08-442-043A-2	Sequence 2, Appl	486	6	2.0	589	2	US-08-453-848-13	Sequence 13, Appl
414	6	2.0	398	2	US-08-641-038A-2	Sequence 2, Appl	487	6	2.0	591	3	US-09-082-737-2	Sequence 2, Appl
415	6	2.0	398	2	US-09-059-178-2	Sequence 2, Appl	488	6	2.0	593	3	US-08-335-865J-21	Sequence 21, Appl
416	6	2.0	398	2	US-08-642-541-2	Sequence 2, Appl	489	6	2.0	598	4	US-08-272-255-22	Sequence 22, Appl
417	6	2.0	398	4	PCT-US91-03478-2	Sequence 2, Appl	490	6	2.0	598	4	PCT-US95-08585-22	Sequence 22, Appl
418	6	2.0	402	3	US-08-948-997-4	Sequence 4, Appl	491	6	2.0	600	3	US-08-904-871-2	Sequence 2, Appl
419	6	2.0	410	3	US-09-083-521-2	Sequence 2, Appl	492	6	2.0	606	3	US-08-335-865J-20	Sequence 20, Appl
420	6	2.0	414	1	US-07-667-276A-4	Sequence 4, Appl	493	6	2.0	607	2	US-08-209-521-7	Sequence 7, Appl
421	6	2.0	418	1	US-08-261-206A-72	Sequence 72, Appl	494	6	2.0	610	3	US-08-961-810-3	Sequence 3, Appl
422	6	2.0	427	3	US-08-448-722A-4	Sequence 4, Appl	495	6	2.0	607	1	US-08-410-804-1	Sequence 1, Appl
423	6	2.0	433	1	US-07-661-610C-2	Sequence 2, Appl	496	6	2.0	610	1	US-08-259-514-1	Sequence 1, Appl
424	6	2.0	433	2	US-08-883-515-2	Sequence 2, Appl	497	6	2.0	610	2	US-08-879-561-10	Sequence 10, Appl
425	6	2.0	434	2	US-09-008-962-1	Sequence 1, Appl	498	6	2.0	610	2	US-08-858-311-1	Sequence 1, Appl
426	6	2.0	434	2	US-08-675-507-1	Sequence 1, Appl	499	6	2.0	622	1	US-08-356-786-16	Sequence 16, Appl
427	6	2.0	434	3	US-09-213-205-1	Sequence 1, Appl	500	6	2.0	626	2	US-07-938-782A-2	Sequence 2, Appl
428	6	2.0	434	3	US-09-012-072-4	Sequence 4, Appl	501	6	2.0	626	1	US-08-630-524-2	Sequence 2, Appl
429	6	2.0	437	3	US-09-073-569-2	Sequence 2, Appl	502	6	2.0	626	4	PCT-US93-08131-2	Sequence 2, Appl
430	6	2.0	441	2	US-08-491-835-4	Sequence 4, Appl	503	6	2.0	634	1	US-08-339-152A-17	Sequence 17, Appl
431	6	2.0	441	3	US-08-946-092A-4	Sequence 4, Appl	504	6	2.0	649	2	US-08-871-266B-16	Sequence 16, Appl
432	6	2.0	441	4	PCT-US94-00685-4	Sequence 4, Appl	505	6	2.0	649	2	US-09-018-864A-16	Sequence 16, Appl
433	6	2.0	450	2	US-08-861-464-2	Sequence 2, Appl	506	6	2.0	649	3	US-08-871-267B-22	Sequence 22, Appl
434	6	2.0	450	2	US-08-665-037-2	Sequence 2, Appl	507	6	2.0	651	2	US-08-431-080-24	Sequence 24, Appl
435	6	2.0	450	2	US-08-396-001-2	Sequence 2, Appl	508	6	2.0	651	2	US-08-938-534-24	Sequence 24, Appl
436	6	2.0	450	2	US-08-666-067-2	Sequence 2, Appl	509	6	2.0	652	1	US-08-261-663A-6	Sequence 6, Appl
437	6	2.0	450	2	US-08-732-870-2	Sequence 2, Appl	510	6	2.0	652	4	PCT-US95-07754A-6	Sequence 6, Appl
438	6	2.0	470	3	US-08-118-319-8	Sequence 8, Appl	511	6	2.0	653	1	US-08-339-152A-16	Sequence 16, Appl
439	6	2.0	471	3	US-08-961-083-158	Sequence 158, App	512	6	2.0	653	2	US-08-007-999B-3	Sequence 3, Appl
440	6	2.0	475	2	US-08-484-200-2	Sequence 2, Appl	513	6	2.0	653	2	US-08-689-276A-3	Sequence 3, Appl
441	6	2.0	475	2	US-08-861-464-14	Sequence 14, Appl	514	6	2.0	675	4	PCT-US95-05008-4	Sequence 4, Appl
442	6	2.0	475	2	US-08-396-001-14	Sequence 14, Appl	515	6	2.0	696	3	US-08-899-437-23	Sequence 23, Appl
443	6	2.0	475	3	US-08-465-375-2	Sequence 2, Appl	516	6	2.0	709	1	US-07-814-964-7	Sequence 7, Appl
444	6	2.0	479	1	US-08-313-553-7	Sequence 7, Appl	517	6	2.0	709	1	US-08-258-442-7	Sequence 7, Appl
445	6	2.0	479	3	US-08-767-993-7	Sequence 7, Appl	518	6	2.0	709	1	US-08-328-809-2	Sequence 2, Appl
446	6	2.0	480	2	US-08-425-989B-11	Sequence 11, Appl	519	6	2.0	709	3	US-09-015-003-2	Sequence 2, Appl
447	6	2.0	481	2	US-08-477-451-19	Sequence 19, Appl	520	6	2.0	709	4	PCT-US92-11107-7	Sequence 7, Appl
448	6	2.0	489	1	US-07-903-103-4	Sequence 4, Appl	521	6	2.0	712	1	US-08-587-889-2	Sequence 2, Appl
449	6	2.0	489	1	US-08-044-619A-4	Sequence 4, Appl	522	6	2.0	712	2	US-08-980-060-5	Sequence 5, Appl
450	6	2.0	489	1	US-08-283-911-4	Sequence 4, Appl	523	6	2.0	712	4	PCT-US96-091937-2	Sequence 2, Appl
451	6	2.0	489	1	US-08-245-500A-5	Sequence 5, Appl	524	6	2.0	720	3	US-08-899-437-6	Sequence 6, Appl
452	6	2.0	489	1	US-08-390-546-5	Sequence 5, Appl	525	6	2.0	753	1	US-08-712-241-2	Sequence 2, Appl
453	6	2.0	489	1	US-08-390-479A-5	Sequence 5, Appl	526	6	2.0	753	1	US-08-712-241-3	Sequence 3, Appl
454	6	2.0	489	1	US-08-557-393-5	Sequence 5, Appl	527	6	2.0	753	4	PCT-US92-10621-3	Sequence 3, Appl
455	6	2.0	489	1	US-08-390-516C-5	Sequence 5, Appl	528	6	2.0	753	4	PCT-US94-02233-3	Sequence 3, Appl
456	6	2.0	489	1	US-08-390-517A-5	Sequence 5, Appl	529	6	2.0	776	1	US-08-198-446B-17	Sequence 17, Appl
457	6	2.0	489	1	US-08-390-515A-5	Sequence 5, Appl	530	6	2.0	776	2	US-08-870-693-17	Sequence 17, Appl
458	6	2.0	489	2	US-08-801-718-5	Sequence 5, Appl	531	6	2.0	777	2	US-08-874-678-3	Sequence 3, Appl
459	6	2.0	489	2	US-08-224-482-2	Sequence 2, Appl	532	6	2.0	777	3	US-08-643-839-3	Sequence 3, Appl
460	6	2.0	497	1	US-08-075-193-4	Sequence 4, Appl	533	6	2.0	783	1	US-08-256-938-2	Sequence 2, Appl
461	6	2.0	497	2	US-08-564-090A-4	Sequence 4, Appl	534	6	2.0	787	1	US-08-256-938-4	Sequence 4, Appl
462	6	2.0	497	3	PCT-US94-06698-4	Sequence 2, Appl	535	6	2.0	787	2	US-08-797-689-16	Sequence 16, Appl
463	6	2.0	497	4	PCT-US94-01101-2	Sequence 2, Appl	536	6	2.0	789	3	US-08-727-308-1	Sequence 1, Appl
464	6	2.0	509	4	US-08-665-926-8	Sequence 8, Appl	537	6	2.0	809	4	PCT-US91-01736-3	Sequence 3, Appl
465	6	2.0	531	2	US-08-789-078-3	Sequence 3, Appl	538	6	2.0	816	1	US-08-190-802A-54	Sequence 54, Appl
466	6	2.0					539	6	2.0	821	1	US-08-198-446B-6	Sequence 6, Appl

540	6	2.0	821	2	US-08-870-693-6	Sequence 6, Appli	613	6	2.0	1363	1	US-08-340-011-4	Sequence 4, Appli
541	6	2.0	829	1	US-08-346-455B-34	Sequence 34, Appl	614	6	2.0	1363	2	US-08-874-678-32	Sequence 32, Appl
542	6	2.0	829	3	US-08-977-221-34	Sequence 34, Appl	615	6	2.0	1363	3	US-08-643-839-32	Sequence 32, Appl
543	6	2.0	829	4	PCT-US95-06613-34	Sequence 34, Appl	616	6	2.0	1363	3	US-08-901-710-4	Sequence 4, Appli
544	6	2.0	870	2	US-08-732-192A-2	Sequence 2, Appli	617	6	2.0	1367	2	US-08-249-867C-2	Sequence 2, Appli
545	6	2.0	874	3	US-08-804-439A-15	Sequence 15, Appl	618	6	2.0	1367	2	US-08-625-819-2	Sequence 2, Appli
546	6	2.0	874	3	US-08-720-229-15	Sequence 15, Appl	619	6	2.0	1367	2	US-08-746-559A-2	Sequence 2, Appli
547	6	2.0	878	3	US-08-941-936-2	Sequence 2, Appli	620	6	2.0	1368	2	US-08-874-678-34	Sequence 34, Appl
548	6	2.0	899	1	US-08-365-689-2	Sequence 2, Appli	621	6	2.0	1368	3	US-08-643-839-34	Sequence 34, Appl
549	6	2.0	899	1	US-08-145-138A-2	Sequence 2, Appli	622	6	2.0	1375	3	US-08-665-259-26	Sequence 26, Appl
550	6	2.0	903	2	US-08-209-521-24	Sequence 24, Appl	623	6	2.0	1375	3	US-08-762-500-36	Sequence 36, Appl
551	6	2.0	903	3	US-08-961-810-134	Sequence 134, App	624	6	2.0	1410	3	US-09-335-409-3	Sequence 3, Appli
552	6	2.0	915	1	US-08-346-455B-69	Sequence 69, Appl	625	6	2.0	1437	3	US-09-061-400-2	Sequence 2, Appli
553	6	2.0	915	4	US-08-977-221-69	Sequence 69, Appl	626	6	2.0	1453	2	US-09-001-273-2	Sequence 2, Appli
554	6	2.0	915	4	PCT-US95-06613-69	Sequence 69, Appl	627	6	2.0	1453	3	US-08-843-459A-2	Sequence 2, Appli
555	6	2.0	933	1	US-07-747-781-2	Sequence 2, Appli	628	6	2.0	1503	3	US-08-976-255-14	Sequence 14, Appl
556	6	2.0	933	4	PCT-US92-06888-2	Sequence 2, Appli	629	6	2.0	1706	2	US-08-459-568-2	Sequence 2, Appli
557	6	2.0	970	2	US-08-673-789-7	Sequence 7, Appli	630	6	2.0	1706	2	US-08-399-411-2	Sequence 2, Appli
558	6	2.0	973	1	US-08-162-809-10	Sequence 10, Appl	631	6	2.0	1706	3	US-08-516-859A-2	Sequence 2, Appli
559	6	2.0	983	2	US-08-164-292B-26	Sequence 26, Appl	632	6	2.0	1996	2	US-08-804-227C-9	Sequence 9, Appli
560	6	2.0	983	3	US-08-845-623-26	Sequence 26, Appl	633	6	2.0	1996	2	US-08-804-198-3	Sequence 3, Appli
561	6	2.0	983	3	US-08-815-927-26	Sequence 26, Appl	634	6	2.0	2237	1	US-08-455-543A-48	Sequence 48, Appl
562	6	2.0	984	1	US-08-257-073-3	Sequence 3, Appli	635	6	2.0	2237	2	US-08-223-305C-48	Sequence 48, Appl
563	6	2.0	984	2	US-08-184-009-120	Sequence 120, App	636	6	2.0	2337	3	US-08-713-118-2	Sequence 2, Appli
564	6	2.0	984	2	US-08-458-356-120	Sequence 120, App	637	6	2.0	2337	3	US-09-452-007-2	Sequence 2, Appli
565	6	2.0	985	4	PCT-US96-03916-6	Sequence 6, Appli	638	6	2.0	2339	1	US-08-455-543A-47	Sequence 47, Appl
566	6	2.0	985	4	PCT-US96-03916-66	Sequence 66, Appl	639	6	2.0	2339	2	US-08-223-305C-47	Sequence 47, Appl
567	6	2.0	988	1	US-08-162-809-14	Sequence 14, Appl	640	6	2.0	2465	2	US-08-596-291-3	Sequence 3, Appli
568	6	2.0	993	1	US-08-348-143-1	Sequence 1, Appli	641	6	2.0	2465	3	US-09-100-804-3	Sequence 3, Appli
569	6	2.0	993	1	US-08-571-785-1	Sequence 1, Appli	642	6	2.0	2466	3	US-09-080-855-12	Sequence 12, Appl
570	6	2.0	998	2	US-08-449-645A-20	Sequence 20, Appl	643	6	2.0	2466	4	PCT-US94-09943-2	Sequence 2, Appli
571	6	2.0	998	2	US-08-702-367A-20	Sequence 20, Appl	644	6	2.0	2485	4	PCT-US94-00198-1	Sequence 1, Appli
572	6	2.0	1003	4	PCT-US95-04681-20	Sequence 20, Appl	645	6	2.0	2485	4	PCT-US94-00198-2	Sequence 2, Appli
573	6	2.0	1003	1	US-07-991-967B-6	Sequence 6, Appli	646	6	2.0	2512	2	US-08-801-263A-9	Sequence 9, Appli
574	6	2.0	1003	1	US-08-107-755A-6	Sequence 6, Appli	647	6	2.0	2512	3	US-09-102-248-9	Sequence 9, Appli
575	6	2.0	1003	2	US-08-544-332-6	Sequence 6, Appli	648	6	2.0	2588	2	US-08-936-135-2	Sequence 2, Appli
576	6	2.0	1007	3	US-08-961-083-216	Sequence 216, App	649	6	2.0	2647	2	US-08-583-562B-8	Sequence 8, Appli
577	6	2.0	1091	3	US-08-986-485-5	Sequence 5, Appli	650	6	2.0	2647	2	US-08-779-113-8	Sequence 8, Appli
578	6	2.0	1101	3	US-08-986-485-2	Sequence 2, Appli	651	6	2.0	2703	1	US-08-185-432-19	Sequence 19, Appl
579	6	2.0	1129	3	US-08-904-871-11	Sequence 11, Appl	652	6	2.0	2818	1	US-08-510-284-1	Sequence 1, Appli
580	6	2.0	1143	2	US-08-310-912A-108	Sequence 108, App	653	6	2.0	2818	1	US-08-411-389-2	Sequence 2, Appli
581	6	2.0	1143	4	PCT-US95-04589-108	Sequence 108, App	654	6	2.0	2818	2	US-08-449-933-2	Sequence 2, Appli
582	6	2.0	1144	1	US-08-261-663A-2	Sequence 2, Appli	655	6	2.0	3033	1	US-07-925-695-8	Sequence 8, Appli
583	6	2.0	1144	1	US-08-261-663A-4	Sequence 4, Appli	656	6	2.0	3033	1	US-07-925-695-9	Sequence 9, Appli
584	6	2.0	1144	3	US-08-930-996A-9	Sequence 9, Appli	657	6	2.0	3218	1	US-08-764-100-27	Sequence 27, Appl
585	6	2.0	1144	3	PCT-US95-07754A-9	Sequence 9, Appli	658	6	2.0	3672	2	US-08-822-445-12	Sequence 12, Appl
586	6	2.0	1144	4	PCT-US95-07754A-2	Sequence 2, Appli	659	6	2.0	3729	2	US-08-804-227C-4	Sequence 4, Appli
587	6	2.0	1144	4	PCT-US95-07754A-4	Sequence 4, Appli	660	6	2.0	3739	3	US-09-320-878-2	Sequence 2, Appli
588	6	2.0	1184	1	US-08-446-038B-20	Sequence 20, Appl	661	6	2.0	3801	2	US-08-822-445-10	Sequence 10, Appl
589	6	2.0	1184	1	US-08-446-010B-20	Sequence 20, Appl	662	6	2.0	4544	1	US-08-469-486-52	Sequence 52, Appl
590	6	2.0	1184	2	US-08-805-445-20	Sequence 20, Appl	663	6	2.0	4544	2	US-08-469-658-52	Sequence 52, Appl
591	6	2.0	1184	2	US-08-064-067D-20	Sequence 20, Appl							
592	6	2.0	1184	2	US-08-066-208-20	Sequence 20, Appl							
593	6	2.0	1187	1	US-08-357-598-8	Sequence 8, Appli							
594	6	2.0	1187	1	US-08-097-997A-13	Sequence 13, Appl							
595	6	2.0	1187	2	US-09-003-289-8	Sequence 8, Appli							
596	6	2.0	1187	3	US-08-665-574C-13	Sequence 13, Appl							
597	6	2.0	1187	4	PCT-US95-16435-8	Sequence 8, Appli							
598	6	2.0	1237	1	US-08-241-853-2	Sequence 2, Appli							
599	6	2.0	1237	2	US-08-850-917-2	Sequence 2, Appli							
600	6	2.0	1243	2	US-08-557-139-2	Sequence 2, Appli							
601	6	2.0	1248	3	US-08-882-046-6	Sequence 6, Appli							
602	6	2.0	1261	1	US-08-764-100-26	Sequence 6, Appli							
603	6	2.0	1288	3	US-08-762-428A-6	Sequence 6, Appli							
604	6	2.0	1294	2	US-08-819-288-3	Sequence 3, Appli							
605	6	2.0	1296	3	US-08-728-603-15	Sequence 15, Appl							
606	6	2.0	1298	1	US-08-222-616-33	Sequence 33, Appl							
607	6	2.0	1298	1	US-08-340-011-2	Sequence 2, Appli							
608	6	2.0	1298	2	US-08-901-710-2	Sequence 2, Appli							
609	6	2.0	1298	4	PCT-US95-04228-33	Sequence 33, Appl							
610	6	2.0	1321	1	US-08-261-822A-3	Sequence 3, Appli							
611	6	2.0	1321	4	PCT-US95-07744A-3	Sequence 3, Appli							
612	6	2.0	1362	2	US-08-874-678-33	Sequence 33, Appl							
	6	2.0	1362	3	US-08-643-839-33	Sequence 33, Appl							

ALIGNMENTS

RESULT 1
US-08-794-796-2
; Sequence 2, Application US/08794796
; Patent No. 5885800
; GENERAL INFORMATION:
; APPLICANT: Emery, John
; APPLICANT: Tan, KB
; APPLICANT: Truneh, Alem
; APPLICANT: Young, Peter
; TITLE OF INVENTION: Tumor Necrosis Related Receptor,
; TITLE OF INVENTION: TR4
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA

```

;
; COUNTRY: USA
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/794,796
; FILING DATE: 04-FEB-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: GH50000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-4026
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 300 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-794-796-2

Query Match 100.0%; Score 300; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 1.4e-267;
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALEGFGLSLCLVLPAALPVPVAVRGVAETPTYPWRDAETGERLVCAQCPTGTFVQR 60
   |||||||
Db 1 MRALEGFGLSLCLVLPAALPVPVAVRGVAETPTYPWRDAETGERLVCAQCPTGTFVQR 60

QY 61 PCRRDSTTCGCPGPRHYTOFWNYLERCYCNVLCGEREEAACHATHNRACRCRTGFF 120
   |||||||
Db 61 PCRRDSTTCGCPGPRHYTOFWNYLERCYCNVLCGEREEAACHATHNRACRCRTGFF 120

QY 121 AHAGFCLHEACPPGAGVAPGTPSONTQCPCPGTFSASSSSSECCOPHRNCTALGLA 180
   |||||||
Db 121 AHAGFCLHEACPPGAGVAPGTPSONTQCPCPGTFSASSSSSECCOPHRNCTALGLA 180

QY 181 LNVPGSSSHDTLCTSGTGFPLSTRVPGAEECERAVIDFVAFQDISIKRLQRLQALEAPE 240
   |||||||
Db 181 LNVPGSSSHDTLCTSGTGFPLSTRVPGAEECERAVIDFVAFQDISIKRLQRLQALEAPE 240

QY 241 GWGPTPRAGAAQLKLRRRLTELLGQDQDGLLVRLQALRVARMPCGLERSVREPLPVH 300
   |||||||
Db 241 GWGPTPRAGAAQLKLRRRLTELLGQDQDGLLVRLQALRVARMPCGLERSVREPLPVH 300

RESULT 2
US-08-938-548B-4
; Sequence 4, Application US/08938548B
; Patent No. 6001963
; GENERAL INFORMATION:
; APPLICANT: Yanagisawa, Masashi
; APPLICANT: Bergsma, Derk
; APPLICANT: Wilson, Shelagh
; APPLICANT: Brooks, David
; APPLICANT: Gellai, Miklos
; TITLE OF INVENTION: NOVEL LIGANDS OF THE NEUROPEPTIDE
; TITLE OF INVENTION: RECEPTOR HFGAN72
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
```

```

;
; CITY: King of Prussia
; STATE: PA
; COUNTRY: United States of America
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,548B
; FILING DATE: 26-SEPT-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/887,382
; FILING DATE: 2-JUL-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/820,519
; FILING DATE: 19-MAR-1997
; APPLICATION NUMBER: 60/033,604
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Elizabeth J. Hecht
; REGISTRATION NUMBER: 41,824
; REFERENCE/DOCKET NUMBER: ATG50037-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5009
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-938-548B-4

Query Match 2.7%; Score 8; DB 3; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLLQA 235
   |||||||
Db 10 RLQRLLQA 17

RESULT 3
US-08-938-548B-9
; Sequence 9, Application US/08938548B
; Patent No. 6001963
; GENERAL INFORMATION:
; APPLICANT: Yanagisawa, Masashi
; APPLICANT: Bergsma, Derk
; APPLICANT: Wilson, Shelagh
; APPLICANT: Brooks, David
; APPLICANT: Gellai, Miklos
; TITLE OF INVENTION: NOVEL LIGANDS OF THE NEUROPEPTIDE
; TITLE OF INVENTION: RECEPTOR HFGAN72
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: United States of America
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
```

```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,548B
; FILING DATE: 26-SEPT-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/887,382
; FILING DATE: 2-JUL-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/820,519
; FILING DATE: 19-MAR-1997
; APPLICATION NUMBER: 60/033,604
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Elizabeth J. Hecht
; REGISTRATION NUMBER: 41,824
; REFERENCE/DOCKET NUMBER: ATG50037-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5009
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-938-548B-9
;
; Query Match 2.7%; Score 8; DB 3; Length 28;
; Best Local Similarity 100.0%; Pred. No. 1.1;
; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
QY 228 RLQRLQA 235
Db 10 RLQRLQA 17
;
RESULT 4
US-08-938-548B-12
; Sequence 12, Application US/08938548B
; Patent No. 6001963
; GENERAL INFORMATION:
; APPLICANT: Yanagisawa, Masashi
; APPLICANT: Bergsma, Derk
; APPLICANT: Wilson, Shelagh
; APPLICANT: Brooks, David
; APPLICANT: Gellai, Miklos
; TITLE OF INVENTION: NOVEL LIGANDS OF THE NEUROPEPTIDE
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: United States of America
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,548B
; FILING DATE: 26-SEPT-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/887,382
; FILING DATE: 2-JUL-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/820,519
```

```
;
; FILING DATE: 19-MAR-1997
; APPLICATION NUMBER: 60/033,604
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Elizabeth J. Hecht
; REGISTRATION NUMBER: 41,824
; REFERENCE/DOCKET NUMBER: ATG50037-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5009
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-938-548B-12
;
; Query Match 2.7%; Score 8; DB 3; Length 28;
; Best Local Similarity 100.0%; Pred. No. 1.1;
; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
QY 228 RLQRLQA 235
Db 10 RLQRLQA 17
;
RESULT 5
US-08-938-548B-10
; Sequence 10, Application US/08938548B
; Patent No. 6001963
; GENERAL INFORMATION:
; APPLICANT: Yanagisawa, Masashi
; APPLICANT: Bergsma, Derk
; APPLICANT: Wilson, Shelagh
; APPLICANT: Brooks, David
; APPLICANT: Gellai, Miklos
; TITLE OF INVENTION: NOVEL LIGANDS OF THE NEUROPEPTIDE
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: United States of America
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,548B
; FILING DATE: 26-SEPT-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/887,382
; FILING DATE: 2-JUL-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/820,519
; FILING DATE: 19-MAR-1997
; APPLICATION NUMBER: 60/033,604
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Elizabeth J. Hecht
; REGISTRATION NUMBER: 41,824
; REFERENCE/DOCKET NUMBER: ATG50037-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5009
```

```

; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 123 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-938-548B-10

Query Match 2.7%; Score 8; DB 3; Length 123;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLLOA 235
Db 71 RLQRLLOA 78

RESULT 6
US-08-938-548B-6
; Sequence 6, Application US/08938548B
; Patent No. 6001963
; GENERAL INFORMATION:
; APPLICANT: Yanagisawa, Masashi
; APPLICANT: Bergsma, Derk
; APPLICANT: Wilson, Shelagh
; APPLICANT: Brooks, David
; APPLICANT: Gellai, Miklos
; TITLE OF INVENTION: NOVEL LIGANDS OF THE NEUROPEPTIDE
; TITLE OF INVENTION: RECEPTOR HFGAN72
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: United States of America
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,548B
; FILING DATE: 26-SEPT-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/887,382
; FILING DATE: 2-JUL-1997
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 130 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-938-548B-10
```

```

US-08-938-548B-6

Query Match 2.7%; Score 8; DB 3; Length 130;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLLOA 235
Db 78 RLQRLLOA 85

RESULT 7
US-08-938-548B-2
; Sequence 2, Application US/08938548B
; Patent No. 6001963
; GENERAL INFORMATION:
; APPLICANT: Yanagisawa, Masashi
; APPLICANT: Bergsma, Derk
; APPLICANT: Wilson, Shelagh
; APPLICANT: Brooks, David
; APPLICANT: Gellai, Miklos
; TITLE OF INVENTION: NOVEL LIGANDS OF THE NEUROPEPTIDE
; TITLE OF INVENTION: RECEPTOR HFGAN72
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: United States of America
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,548B
; FILING DATE: 26-SEPT-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/887,382
; FILING DATE: 2-JUL-1997
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 131 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-938-548B-2

Query Match 2.7%; Score 8; DB 3; Length 131;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLLOA 235
Db 78 RLQRLLOA 85
```

Db 79 RLQRLQA 86

RESULT 8

US-08-313-288B-19
; Sequence 19, Application US/08313288B
; Patent No. 5750502
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M. and Avihu Klar
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/313,288B
; FILING DATE: January 5, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0526
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1172 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-313-288B-19

Query Match 2.7%; Score 8; DB 1; Length 1172;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 ALEGPGLS 10

Db 102 ALEGPGLS 109

RESULT 9

US-08-804-227C-2
; Sequence 2, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: DeHoff, Bradley S.
; APPLICANT: Kuhstoss, Stuart A.
; APPLICANT: Rostock, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCI(DOS) Text only
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,227C
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas, G.
; REGISTRATION NUMBER: 35,784
; REFERENCE/DOCKET NUMBER: X-8231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4472 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-804-227C-2

Query Match 2.7%; Score 8; DB 2; Length 4472;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 257 LRRRLTEL 264

Db 4424 LRRRLTEL 4431

RESULT 10

US-07-776-272-25
; Sequence 25, Application US/07776272
; Patent No. 5612454
; GENERAL INFORMATION:
; APPLICANT: Kaminuma, Toshihiko
; APPLICANT: Iida, Toshii
; APPLICANT: Tajima, Masahiro
; TITLE OF INVENTION: Process for Purification of Polypeptide
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wegner, Cantor, Mueller & Player
; STREET: 1233 20th St. N.W. P.O. Box 18218
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: United States of America
; ZIP: 20036-8218
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/776,272
; FILING DATE: 19911129
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Player, William E.
; REGISTRATION NUMBER: 31,409
; REFERENCE/DOCKET NUMBER: P-450-23167
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-887-0400
; TELEFAX: 202-887-0605
; TELEX: 440706
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

```
; HYPOTHETICAL: YES
US-07-776-272-25

Query Match          2.3%; Score 7; DB 1; Length 26;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLQ 234
Db 17 RLQRLQ 23

RESULT 11
US-07-924-054-10
; Sequence 10, Application US/07924054
; Patent No. 5486472
; GENERAL INFORMATION:
; APPLICANT: SUZUKI, No. 5486472uhiro
; APPLICANT: TSUDA, Masao
; TITLE OF INVENTION: ANTIBODY TO PACAP AND USE THEREOF
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS&
; ADDRESSEE: CUSHMAN
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 19920903
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RESNICK, David S
; REGISTRATION NUMBER: 34235
; REFERENCE/DOCKET NUMBER: 40805
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)523-3400
; TELEFAX: (617)523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-924-054-10

Query Match          2.3%; Score 7; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLQ 234
Db 18 RLQRLQ 24

RESULT 12
US-08-062-472B-43
; Sequence 43, Application US/08062472B
; Patent No. 5695954
; GENERAL INFORMATION:
; APPLICANT: Sherwood, Nancy G M
; APPLICANT: Parker, David B

; APPLICANT: MCROITY, John E
; APPLICANT: Lescheid, David W
; TITLE OF INVENTION: DNA ENCODING TWO FISH NEUROPEPTIDES
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KIARQUIST, SPARKMAN, CAMPBELL, LEIGH &
; ADDRESSEE: WHINSTON, LLP
; STREET: ONE WORLD TRADE CENTER, SUITE 1600, 121 S.W.
; STREET: SALMON STREET
; CITY: PORTLAND
; STATE: OREGON
; COUNTRY: USA
; ZIP: 97204-2988
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/062.472B
; FILING DATE: 14-MAY-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: POLLEY, RICHARD J
; REGISTRATION NUMBER: 28107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (503) 226-7391
; TELEFAX: (503) 228-9446
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-062-472B-43

Query Match          2.3%; Score 7; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLQRLQ 234
Db 18 RLQRLQ 24

RESULT 13
US-08-519-180-6
; Sequence 6, Application US/08519180
; Patent No. 5770570
; GENERAL INFORMATION:
; APPLICANT: PAUL, SUDHIR
; APPLICANT: YASUKO, NODA
; APPLICANT: ISRAEL, RUBINSTEIN
; TITLE OF INVENTION: A METHOD OF DELIVERING A VASOACTIVE
; TITLE OF INVENTION: INTESTINAL POLYPEPTIDE, AN ENCAPSULATED VASOACTIVE
; TITLE OF INVENTION: INTESTINAL POLYPEPTIDE, AND A METHOD OF MAKING THE
; TITLE OF INVENTION: ENCAPSULATED VASOACTIVE INTESTINAL POLYPEPTIDE
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

; APPLICATION NUMBER: US/08/519,180
; FILING DATE: 25-AUG-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION NUMBER: US 08/224488
; FILING DATE: 07-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: SEMINAUER, JEFFREY A.
; REGISTRATION NUMBER: 31,933
; REFERENCE/DOCKET NUMBER: 4464/98971
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-519-180-6

Query Match 2.3%; Score 7; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLRLQLQ 234
Db 18 RLRLQLQ 24

RESULT 14
US-08-818-253-36
; Sequence 36, Application US/08818253
; Patent No. 5998204
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Miyawaki, Atsushi
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; TITLE OF INVENTION: DETECTION OF ANALYTES
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/818,253
; FILING DATE: 14-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07257/043001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; MOLECULE TYPE: peptide
US-08-818-253-36

Query Match 2.3%; Score 7; DB 2; Length 27;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 RLRLQLQ 234
Db 18 RLRLQLQ 24

RESULT 15
US-08-583-569-1
; Sequence 1, Application US/08583569
; Patent No. 5869234
; GENERAL INFORMATION:
; APPLICANT: Knipe, David M.
; APPLICANT: Xia, Kai
; APPLICANT: DeLuca, Neal A.
; TITLE OF INVENTION: METHOD OF IDENTIFYING COMPOUNDS
; TITLE OF INVENTION: WHICH
; TITLE OF INVENTION: MODULATE HERPESVIRUS INFECTION
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,569
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: HU95-08
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 69 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-583-569-1

Query Match 2.3%; Score 7; DB 2; Length 69;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
Db 49 SASSSSS 55

RESULT 16
US-08-374-843B-6
; Sequence 6, Application US/08374843B
; Patent No. 5726016
; GENERAL INFORMATION:
; APPLICANT: Demuth, Donald R.
; APPLICANT: Lally, Edward T.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
TITLE OF INVENTION: OF DISEASES ASSOCIATED WITH ACTINOBACILLUS
TITLE OF INVENTION: ACTINOMYCETEMCOMITANS INFECTION
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
STREET: 1601 Market Street, 36th Street
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19103-2398
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/374,843B
FILING DATE: 18-JAN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Leary, Kathryn Ph.D.
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: 9596-7 (F-1080)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 567-2020
TELEFAX: (215) 567-2991
TELEX: 831-494
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 150 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-374-843B-6

Query Match 2.3%; Score 7; DB 1; Length 150;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 LRRRLTE 263
|||||
Db 133 LRRRLTE 139

RESULT 17
US-08-374-843B-10
Sequence 10, Application US/08374843B
Patent No. 5726016
GENERAL INFORMATION:
APPLICANT: DeMuth, Donald R.
APPLICANT: Lally, Edward T.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
TITLE OF INVENTION: OF DISEASES ASSOCIATED WITH ACTINOBACILLUS
TITLE OF INVENTION: ACTINOMYCETEMCOMITANS INFECTION
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
STREET: 1601 Market Street, 36th Street
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19103-2398
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/374,843B
FILING DATE: 18-JAN-1995

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Leary, Kathryn Ph.D.
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: 9596-7 (F-1080)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 567-2020
TELEFAX: (215) 567-2991
TELEX: 831-494
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 150 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-374-843B-10

Query Match 2.3%; Score 7; DB 1; Length 150;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 LRRRLTE 263
|||||
Db 133 LRRRLTE 139

RESULT 18
US-08-905-420-6
Sequence 6, Application US/08905420
Patent No. 5861255
GENERAL INFORMATION:
APPLICANT: DeMuth, Donald R.
APPLICANT: Lally, Edward T.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
TITLE OF INVENTION: OF DISEASES ASSOCIATED WITH ACTINOBACILLUS
TITLE OF INVENTION: ACTINOMYCETEMCOMITANS INFECTION
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
STREET: 1601 Market Street, 36th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19103-2398
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/905,420
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/374,843
FILING DATE: 18-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Leary, Kathryn Ph.D.
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: 9596-7 (F-1080)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 567-2020
TELEFAX: (215) 567-2991
TELEX: 831-494
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 150 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-905-420-6

Query Match 2.3%; Score 7; DB 2; Length 150;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 LRRRLTE 263
|||||
DB 133 LRRRLTE 139

RESULT 19

US-08-905-420-10
; Sequence 10, Application US/08905420
; Patent No. 5861255
; GENERAL INFORMATION:
; APPLICANT: DeMuth, Donald R.
; APPLICANT: Lally, Edward T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSIS
; TITLE OF INVENTION: OF DISEASES ASSOCIATED WITH ACTINOBACILLUS
; TITLE OF INVENTION: ACTINOMYCETEMCOMITANS INFECTION
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
; STREET: 1601 Market Street, 36th Floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103-2398

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: US/08/905.420
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/374,843
FILING DATE: 18-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Leary, Kathryn Ph.D.
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: 9596-7 (F-1080)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 567-2020
TELEFAX: (215) 567-2991
TELEX: 831-494
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 150 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-905-420-10

Query Match 2.3%; Score 7; DB 2; Length 150;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 LRRRLTE 263
|||||
DB 133 LRRRLTE 139

RESULT 20

US-08-185-432-6
; Sequence 6, Application US/08185432
; Patent No. 5750652
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, Spyridon

; APPLICANT: Busseau, Isabelle
; APPLICANT: Diederich, Robert J.
; APPLICANT: Xu, Tian
; APPLICANT: Matsuno, Kenji
; TITLE OF INVENTION: DELTEX PROTEINS, NUCLEIC ACIDS, AND
; TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185.432
; FILING DATE: 21-JAN-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 181 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-185-432-6

Query Match 2.3%; Score 7; DB 1; Length 181;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
|||||
DB 40 SASSSSS 46

RESULT 21

US-09-066-408-12
; Sequence 12, Application US/09066408
; Patent No. 6060448
; GENERAL INFORMATION:
; APPLICANT: Smith, John Arthur
; APPLICANT: Wilkinson, Mark Charles
; APPLICANT: Liu, Qing-Ming
; TITLE OF INVENTION: Casein Fragments Having Growth Promoting
; TITLE OF INVENTION: Activity
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```
; APPLICATION NUMBER: US/09/066,408
; FILING DATE: 13-MAR-1998
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB96/02658
; FILING DATE: 31-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9522302.0
; FILING DATE: 31-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 018317-000100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 235 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..235
; OTHER INFORMATION: /note= "pig alpha-S2 casein
; OTHER INFORMATION: precursor"
;
US-09-066-408-12

Query Match                2.3%; Score 7; DB 3; Length 235;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 ASSSSSE 166
   |||||
Db 69 ASSSSSE 75

RESULT 22
PCT-US96-12374-2
; Sequence 2, Application PC/TUS9612374
; GENERAL INFORMATION:
; APPLICANT: Northwestern University
; TITLE OF INVENTION: Herpes Virus Entry Mediator
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Goldsmith, Milnamow & Katz, Ltd.
; STREET: 180 N. Stetson, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/12374
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Northrup, Thomas E.
; REGISTRATION NUMBER: 33,268
; REFERENCE/DOCKET NUMBER: NOR3446P020PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5400
; TELEFAX: (312) 616-5460
; TELEX: --
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 283 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US96-12374-2

Query Match                2.3%; Score 7; DB 4; Length 283;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 153 CPPGTFS 159
   |||||
Db 165 CPPGTFS 171

RESULT 23
US-08-965-903B-8
; Sequence 8, Application US/08965903B
; Patent No. 6060275
; GENERAL INFORMATION:
; APPLICANT: Hachon, Nir
; APPLICANT: Krasnow, Mark A.
; TITLE OF INVENTION: SPROUTY PROTEIN AND CODING
; TITLE OF INVENTION: SEQUENCE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/965,903B
; FILING DATE: 07-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/030232
; FILING DATE: 07-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Petithory, Joanne R
; REGISTRATION NUMBER: 42,995
; REFERENCE/DOCKET NUMBER: 8600-0177.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; TELEX:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 315 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: h-spry2 protein
;
US-08-965-903B-8

Query Match                2.3%; Score 7; DB 3; Length 315;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 161 SSSSSEQ 167
   |||||
Db 127 SSSSSEQ 133
```

RESULT 24
US-08-888-429A-21
; Sequence 21, Application US/08888429A
; Patent No. 6136596
; GENERAL INFORMATION:
; APPLICANT: Davis, Roger J.
; APPLICANT: Whitmarsh, Alan
; APPLICANT: Tournier, Cathy
; TITLE OF INVENTION: CYTOKINE-, STRESS-, AND ONCOPROTEIN-
; TITLE OF INVENTION: ACTIVATED HUMAN PROTEIN KINASE KINASES
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/888,429A
; FILING DATE: 07-JUL-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/530,950
; FILING DATE: 19-SEP-1995
; APPLICATION NUMBER: 08/446,083
; FILING DATE: 19-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, Peter J.
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 07917/053001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 299354
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-888-429A-21

Query Match 2.3%; Score 7; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
Db 1 SASSSSS 7
|||||||

RESULT 25
US-08-123-343A-5
; Sequence 5, Application US/08123343A
; Patent No. 5593879
; GENERAL INFORMATION:
; APPLICANT: Steller, Hermann
; APPLICANT: Abrams, John M.
; APPLICANT: Grether, Megan E.
; APPLICANT: White, Kristin
; TITLE OF INVENTION: Cell Death Genes of Drosophila
; TITLE OF INVENTION: Melanogaster and Vertebrate Analogs
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/123,343A
; FILING DATE: 17-SEP-1993
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/004,957
; FILING DATE: 15-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: MIT-5907A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 61861-9540
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 410 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-123-343A-5

Query Match 2.3%; Score 7; DB 1; Length 410;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
Db 53 SASSSSS 59
|||||||

RESULT 26
US-08-123-343A-7
; Sequence 7, Application US/08123343A
; Patent No. 5593879
; GENERAL INFORMATION:
; APPLICANT: Steller, Hermann
; APPLICANT: Abrams, John M.
; APPLICANT: Grether, Megan E.
; APPLICANT: White, Kristin
; TITLE OF INVENTION: Cell Death Genes of Drosophila
; TITLE OF INVENTION: Melanogaster and Vertebrate Analogs
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/123,343A
; FILING DATE: 17-SEP-1993
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/004,957

;; FILING DATE: 15-JAN-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Granahan, Patricia
;; REGISTRATION NUMBER: 32,227
;; REFERENCE/DOCKET NUMBER: MIT-5907A
;; TELEPHONE: 617-861-6240
;; TELEFAX: 61861-9540
;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 410 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: unknown
;; TOPOLOGY: unknown
;; MOLECULE TYPE: protein
US-08-123-343A-7

Query Match 2.3%; Score 7; DB 1; Length 410;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
Db 53 SASSSSS 59

RESULT 27
US-08-318-947A-16
;; Sequence 16, Application US/08318947A
;; Patent No. 5798245
;; GENERAL INFORMATION:
;; APPLICANT: Anderson, Paul J.
;; APPLICANT: Tian, Qingsheng
;; TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED
;; NUMBER OF SEQUENCES: 21
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
;; STREET: 2100 Pennsylvania Avenue, NW Suite 800
;; CITY: Washington
;; STATE: DC
;; COUNTRY: USA
;; ZIP: 20037
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/318,947A
;; FILING DATE: 06-OCT-1994
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/133,530
;; FILING DATE: 07-OCT-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Mack, Susan J.
;; REGISTRATION NUMBER: 30,951
;; REFERENCE/DOCKET NUMBER: A6462
;; TELEPHONE: (202)293-7060
;; TELEFAX: (202)293-2920
;; TELEX: 6491103
;; INFORMATION FOR SEQ ID NO: 16:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 430 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-318-947A-16

Query Match 2.3%; Score 7; DB 1; Length 430;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 229 LQRLQQA 235
Db 21 LQRLQQA 27
RESULT 28
US-08-795-303-16
;; Sequence 16, Application US/08795303
;; Patent No. 5948656
;; GENERAL INFORMATION:
;; APPLICANT: Anderson, Paul J.
;; APPLICANT: Tian, Qingsheng
;; TITLE OF INVENTION: TIA-1 BINDING PROTEINS AND ISOLATED
;; NUMBER OF SEQUENCES: 21
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
;; STREET: 2100 Pennsylvania Avenue, NW Suite 800
;; CITY: Washington
;; STATE: DC
;; COUNTRY: USA
;; ZIP: 20037
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/795,303
;; FILING DATE: 04-FEB-1997
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/318,947
;; FILING DATE: 06-OCT-1994
;; APPLICATION NUMBER: 08/133,530
;; FILING DATE: 07-OCT-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Mack, Susan J.
;; REGISTRATION NUMBER: 30,951
;; REFERENCE/DOCKET NUMBER: A6462
;; TELEPHONE: (202)293-7060
;; TELEFAX: (202)293-2920
;; TELEX: 6491103
;; INFORMATION FOR SEQ ID NO: 16:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 430 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-795-303-16

Query Match 2.3%; Score 7; DB 2; Length 430;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 229 LQRLQQA 235
Db 21 LQRLQQA 27

RESULT 29
US-08-946-241B-2
;; Sequence 2, Application US/08946241B
;; Patent No. 5928941
;; GENERAL INFORMATION:
;; APPLICANT: Lee, Mu-En
;; APPLICANT: MCA/Nulty, Megan M.

```
;
; TITLE OF INVENTION: REPRESSOR KRUPPEL-LIKE FACTOR
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/946,241B
; FILING DATE: 07-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/030,035
; FILING DATE: 05-NOV-1996
; APPLICATION NUMBER: 60/027,521
; FILING DATE: 07-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Creason, Gary L.
; REGISTRATION NUMBER: 34,310
; REFERENCE/DOCKET NUMBER: 05433/027001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-8906
; TELEFAX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 470 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-946-241B-2

Query Match 2.3%; Score 7; DB 2; Length 470;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 159 SASSSSS 165
Db 120 SASSSSS 126

RESULT 30
US-09-309-053-2
; Sequence 2, Application US/09309053
; Patent No. 6077933
; GENERAL INFORMATION:
; APPLICANT: Lee, Mu-En
; TITLE OF INVENTION: REPRESSOR KRUPPEL-LIKE FACTOR
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/309,053
; FILING DATE: 07-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/030,035
; FILING DATE: 05-NOV-1996
; APPLICATION NUMBER: 60/027,521
; FILING DATE: 07-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Creason, Gary L.
; REGISTRATION NUMBER: 34,310
; REFERENCE/DOCKET NUMBER: 05433/027001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-8906
; TELEFAX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 470 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-946-241B-2
```

```
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/946,241
; FILING DATE: 07-OCT-1997
; APPLICATION NUMBER: 60/030,035
; FILING DATE: 05-NOV-1996
; APPLICATION NUMBER: 60/027,521
; FILING DATE: 07-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Creason, Gary L.
; REGISTRATION NUMBER: 34,310
; REFERENCE/DOCKET NUMBER: 05433/027001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-8906
; TELEFAX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 470 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-09-309-053-2

Query Match 2.3%; Score 7; DB 3; Length 470;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 159 SASSSSS 165
Db 120 SASSSSS 126

RESULT 31
US-08-946-241B-9
; Sequence 9, Application US/08946241B
; Patent No. 5928941
; GENERAL INFORMATION:
; APPLICANT: Lee, Mu-En
; TITLE OF INVENTION: REPRESSOR KRUPPEL-LIKE FACTOR
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/946,241B
; FILING DATE: 07-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/030,035
; FILING DATE: 05-NOV-1996
; APPLICATION NUMBER: 60/027,521
; FILING DATE: 07-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Creason, Gary L.
; REGISTRATION NUMBER: 34,310
; REFERENCE/DOCKET NUMBER: 05433/027001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-8906
; TELEFAX: 200154
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
```

LENGTH: 479 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-946-241B-9

Query Match 2.3%; Score 7; DB 2; Length 479;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
Db 129 SASSSSS 135

RESULT 32
US-09-309-053-9
Sequence 9, Application US/09309053
Patent No. 6077933
GENERAL INFORMATION:
APPLICANT: Lee, Mu-En
APPLICANT: McA Nulty, Megan M.
TITLE OF INVENTION: REPRESSOR KRUPPEL-LIKE FACTOR
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/309,053
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/946,241
FILING DATE: 07-OCT-1997
FILING DATE: 07-OCT-1997
APPLICATION NUMBER: 60/030,035
FILING DATE: 05-NOV-1996
APPLICATION NUMBER: 60/027,521
FILING DATE: 07-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Creason, Gary L.
REGISTRATION NUMBER: 34,310
REFERENCE/DOCKET NUMBER: 05433/027001
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 479 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-09-309-053-9

Query Match 2.3%; Score 7; DB 3; Length 479;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
Db 129 SASSSSS 135

RESULT 33
US-08-469-412A-7
Sequence 7, Application US/08469412A
Patent No. 5856125
GENERAL INFORMATION:
APPLICANT: Mavrothalassitis, George J.
APPLICANT: Blair, Donald G.
APPLICANT: Fisher, Robert J.
APPLICANT: Beal Jr., Gregory J.
APPLICANT: Athanasios, Meropi A.
TITLE OF INVENTION: The ERF Genetic Locus and Its Products
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,412A
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Garrett-Wackowski, Eugenia
REGISTRATION NUMBER: 37,330
REFERENCE/DOCKET NUMBER: 015280-229000
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 543 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..543
OTHER INFORMATION: /note= "murine ERF amino acid sequence
OTHER INFORMATION: (first 8 amino acids from first exon not
included)"
US-08-469-412A-7

Query Match 2.3%; Score 7; DB 2; Length 543;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
Db 355 SASSSSS 361

RESULT 34
US-08-469-412A-2
Sequence 2, Application US/08469412A
Patent No. 5856125
GENERAL INFORMATION:
APPLICANT: Mavrothalassitis, George J.
APPLICANT: Blair, Donald G.
APPLICANT: Fisher, Robert J.
APPLICANT: Beal Jr., Gregory J.
APPLICANT: Athanasios, Meropi A.

; APPLICANT: Sgouras, Dionyssios N.
; TITLE OF INVENTION: The ERF Genetic Locus and Its Products
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,412A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Garrett-Wackowski, Eugenia
; REGISTRATION NUMBER: 37,330
; REFERENCE/DOCKET NUMBER: 015280-229000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 548 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-469-412A-2

Query Match 2.3%; Score 7; DB 2; Length 548;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
DB 363 SASSSSS 369

RESULT 35
US-08-185-432-2
; Sequence 2, Application US/08185432
; Patent No. 5750652
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Busseau, Isabelle
; APPLICANT: Diederich, Robert J.
; APPLICANT: Xu, Tian
; APPLICANT: Matsuno, Kenji
; TITLE OF INVENTION: DELTEX PROTEINS, NUCLEIC ACIDS, AND
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185.432
; FILING DATE: 21-JAN-1994
; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:
; NAME: Mistock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 737 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-185-432-2

Query Match 2.3%; Score 7; DB 1; Length 737;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
DB 345 SASSSSS 351

RESULT 36
US-08-185-432-4
; Sequence 4, Application US/08185432
; Patent No. 5750652
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Busseau, Isabelle
; APPLICANT: Diederich, Robert J.
; APPLICANT: Xu, Tian
; APPLICANT: Matsuno, Kenji
; TITLE OF INVENTION: DELTEX PROTEINS, NUCLEIC ACIDS, AND
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185.432
; FILING DATE: 21-JAN-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 737 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-185-432-4

Query Match 2.3%; Score 7; DB 1; Length 737;

Best Local Similarity 100.0%; Pred. No. 1.7e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

QY 159 SASSSSS 165
| | | | | | | |
DB 345 SASSSSS 351

RESULT 37
US-08-416-950-11
; Sequence 11, Application US/08416950
; Patent No. 5780036
; GENERAL INFORMATION:
; APPLICANT: CHISARI, Francis V.
; TITLE OF INVENTION: PEPTIDES FOR INDUCING CYTOTOXIC T
; TITLE OF INVENTION: LYMPHOCYTE RESPONSES TO HEPATITIS B VIRUS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94105-1492
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/416,950
; FILING DATE: 26-AUG-1992
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US
; FILING DATE: 26-AUG-1992
; FILING DATE: 26-AUG-1992
; FILING DATE: 26-AUG-1992
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14740-2-2
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 845 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-416-950-11

Query Match 2.3%; Score 7; DB 1; Length 845;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0;

QY 159 SASSSSS 165
| | | | | | | |
DB 270 SASSSSS 276

RESULT 38
US-08-469-830-11
; Sequence 11, Application US/08469830
; Patent No. 5932224
; GENERAL INFORMATION:
; APPLICANT: CHISARI, Francis V.

; TITLE OF INVENTION: PEPTIDES FOR INDUCING CYTOTOXIC T
; TITLE OF INVENTION: LYMPHOCYTE RESPONSES TO HEPATITIS B VIRUS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94105-1492
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,830
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/100,870
; FILING DATE: 02-AUG-1993
; APPLICATION NUMBER: US 07/935,898
; FILING DATE: 26-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,540
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14740-2-2
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 845 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-469-830-11

Query Match 2.3%; Score 7; DB 2; Length 845;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0;

QY 159 SASSSSS 165
| | | | | | | |
DB 270 SASSSSS 276

RESULT 39
US-08-785-241-5
; Sequence 5, Application US/08785241
; Patent No. 5695963
; GENERAL INFORMATION:
; APPLICANT: McKnight, Steven L.
; APPLICANT: Russell, David W.
; APPLICANT: Tian, Hui
; TITLE OF INVENTION: Endothelial PAS Domain Protein
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/785.241
FILING DATE: 17-JAN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UTSD:1229
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 875 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-785-241-5

Query Match 2.3%; Score 7; DB 1; Length 875;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 159 SASSSSS 165
Db 472 SASSSSS 478

RESULT 40
US-08-480-662-2
Sequence 2, Application US/08480662
Patent No. 5759782
GENERAL INFORMATION:
APPLICANT: Pastan, Ira
TITLE OF INVENTION: CELLULAR APOPTOSIS SUSCEPTIBILITY PROTEIN (CSP) AND AN
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson and Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Israelson, Ned A
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH12.001A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 971 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
US-08-480-662-2

Query Match 2.3%; Score 7; DB 1; Length 971;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 273 LVRLLQA 279
Db 704 LVRLLQA 710

RESULT 41
US-08-918-190-2
Sequence 2, Application US/08918190
Patent No. 6072031
GENERAL INFORMATION:
APPLICANT: Pastan, Ira
TITLE OF INVENTION: CELLULAR APOPTOSIS SUSCEPTIBILITY
TITLE OF INVENTION: CELLULAR APOPTOSIS SUSCEPTIBILITY
TITLE OF INVENTION: PROTEIN (CSP) AND ANTISENSE CSP
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson and Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: USA
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/480,662
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Israelson, Ned A
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH12.001A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-235-8550
TELEFAX: 619-235-0176
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 971 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
US-08-918-190-2

Query Match 2.3%; Score 7; DB 3; Length 971;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 273 LVRLQA 279
Db 704 LVRLQA 710

RESULT 42

PCT-US96-09927-2
; Sequence 2, Application PC/TUS9609927
; GENERAL INFORMATION:
; APPLICANT: The United States, As Represented by the
; APPLICANT: Secretary, Department of Health and Human
; APPLICANT: Services
; TITLE OF INVENTION: CELLULAR APOPTOSIS
; TITLE OF INVENTION: SUSCEPTIBILITY PROTEIN (CSP) AND ANTISENSE CSP
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson and Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/09927
; FILING DATE: 07-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US08/480662
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH112.001QPC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 971 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
PCT-US96-09927-2

Query Match 2.3%; Score 7; DB 4; Length 971;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 273 LVRLQA 279
Db 704 LVRLQA 710

RESULT 43

US-08-545-860D-55
; Sequence 55, Application US/08545860D
; Patent No. 6040140
; GENERAL INFORMATION:
; APPLICANT: Croce, Carlo
; APPLICANT: Canaan, Eli
; TITLE OF INVENTION: Diagnostics, Therapeutics and Methods

; TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias
; Resulting from Chromosome Abnormalities in the All-1 Region
; NUMBER OF SEQUENCES: 94
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
; ADDRESSEE: No. 6040140ris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,860D
; FILING DATE: 07-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04496
; FILING DATE: 22-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10930
; FILING DATE: 09-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/327,392
; FILING DATE: 19-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/320,559
; FILING DATE: 11-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/062,443
; FILING DATE: 14-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/971,094
; FILING DATE: 30-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/888,839
; FILING DATE: 27-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/805,093
; FILING DATE: 11-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca Esq., Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1262
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1093 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-545-860D-55

Query Match 2.3%; Score 7; DB 3; Length 1093;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
Db 326 SASSSSS 332

RESULT 44

PCT-US94-04496-55
; Sequence 55, Application PC/TUS9404496
; GENERAL INFORMATION:
; APPLICANT: Croce, Carlo
; APPLICANT: Canani, Eli
; TITLE OF INVENTION: Diagnostics, Therapeutics and Methods
; TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias
; TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the All-1
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
; ADDRESSEE: Norris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/04496
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: DeLuca Esq., Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-1242
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1093 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
PCT-US94-04496-55

Query Match 2.3%; Score 7; DB 4; Length 1093;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
Db 326 SASSSSS 332

RESULT 45
US-08-690-473-2
; Sequence 2, Application US/08690473
; Patent No. 5876923
; GENERAL INFORMATION:
; APPLICANT: Leopoldi, Rosario
; APPLICANT: Roizman, Bernard
; TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 AS AN
; TITLE OF INVENTION: INHIBITOR OF APOPTOSIS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/690,473
; FILING DATE: 26-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Highlander, Steven L.
; REGISTRATION NUMBER: 37,642
; REFERENCE/DOCKET NUMBER: ARCD:239
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1298 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-08-690-473-2

Query Match 2.3%; Score 7; DB 2; Length 1298;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
Db 191 SASSSSS 197

RESULT 46
US-08-559-303B-78
; Sequence 78, Application US/08559303B
; Patent No. 5824501
; GENERAL INFORMATION:
; APPLICANT: NATHAN A. ELLIS, JAMES GERMAN, AND JOANNA
; APPLICANT: GRODEN
; TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BLOOM'S SYNDROME
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMSTER, ROTHSTEIN & EBENSTEIN
; STREET: 90 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH 1.44 Mb STORAGE DISKETTE
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/559,303B
; FILING DATE: NOVEMBER 15, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ELIZABETH A. BOGOSIAN
; REGISTRATION NUMBER: 39,911
; REFERENCE/DOCKET NUMBER: 63475/65
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-5995
; TELEFAX: (212) 286-0854 or 286-0082
; TELEX: TWX 710-581-4766
; INFORMATION FOR SEQ ID NO: 78:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1417
; TYPE: AMINO ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE:
; DESCRIPTION: OTHER NUCLEIC ACID

;
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
US-08-559-303B-78

Query Match 2.3%; Score 7; DB 2; Length 1417;
Best Local Similarity 100.0%; Pred. No. 3.1e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
DB 310 SASSSSS 316

RESULT 47
US-08-781-891-78
; Sequence 78, Application US/08781891
; Patent No. 6090620
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Yu, Chang-En
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; APPLICANT: Schellenberg, Gerald D.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; TITLE OF INVENTION: WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781.891
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6090620tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 240052.419
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 78:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1417 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-781-891-78

Query Match 2.3%; Score 7; DB 3; Length 1417;
Best Local Similarity 100.0%; Pred. No. 3.1e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 SASSSSS 165
DB 310 SASSSSS 316

RESULT 48
US-08-804-227C-5
; Sequence 5, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: DeHoff, Bradley S.
; APPLICANT: Kuhstoss, Stuart A.
; APPLICANT: Rostock, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII(DOS) Text only
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,227C
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas G.
; REGISTRATION NUMBER: 35,784
; REFERENCE/DOCKET NUMBER: X-8231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1611 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-804-227C-5

Query Match 2.3%; Score 7; DB 2; Length 1611;
Best Local Similarity 100.0%; Pred. No. 3.5e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 REEEARA 104
DB 1516 REEEARA 1522

RESULT 49
US-08-482-651-10
; Sequence 10, Application US/08482651
; Patent No. 5874409
; GENERAL INFORMATION:
; APPLICANT: Victoria, Edward J.
; APPLICANT: Marquis, David M.
; TITLE OF INVENTION: aPL IMMUNOREACTIVE PEPTIDES, CONJUGATES
; TITLE OF INVENTION: THEREOF AND METHODS OF TREATMENT FOR aPL ANTIBODY-MEDIATED
; TITLE OF INVENTION: PATHOLOGIES
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482.651
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Park, Freddie K.
REGISTRATION NUMBER: 35,636
REFERENCE/DOCKET NUMBER: 25231-20061.00
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: xyz (details pg. 16)
CLONE: 3B6
US-08-482-651-10

Query Match 2.0%; Score 6; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 CLV LAL 18
 |||||
DB 1 CLV LAL 6

RESULT 50
US-08-752-891-7
Sequence 7, Application US/08752891
Patent No. 5837819
GENERAL INFORMATION:
APPLICANT: MATSUMOTO, Kunihiro
APPLICANT: NISHIDA, Eisuke
TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/752.891
FILING DATE: 20-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-300856
FILING DATE: 28-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-126282
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 17981/111
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136

INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-752-891-7

Query Match 2.0%; Score 6; DB 2; Length 24;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 ASSSSS 165
 |||||
DB 7 ASSSSS 12

Search completed: January 30, 2001, 17:07:07
Job time: 977 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 17:06:44 ; Search time 29.47 seconds
(without alignments)
197.250 Million cell updates/sec

Title: US-09-518-931-4

Perfect score: 170

Sequence: 1 MRALEPGLSLLCLLVLPA.....PRSGRRRCRGQVAGPSLAP 170

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 268485 seqs, 34193795 residues

Word size : 6

Total number of hits satisfying chosen parameters: 300

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : A_Geneseq_36.*

1: /SIDS1/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq/AA2000.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	170	100.0	170	19	Human tumour necro
2	142	83.5	245	20	A human tumour nec
3	142	83.5	300	19	Amino acid sequenc
4	142	83.5	300	19	Human tumour necro
5	142	83.5	300	20	Human lung TNF-rec
6	142	83.5	300	20	Human FLINT #1 pro
7	142	83.5	300	20	Mammalian tumour n
8	142	83.5	300	20	Human Dcr3 polypep
9	142	83.5	300	20	Human tumour necro
10	142	83.5	300	20	Orphan receptor (H
11	116	68.2	302	20	Human FLINT #2 pro
12	113	66.5	271	20	Human mFLINT #1 pr

13	87	51.2	273	20	Y42185	Human mFLINT #2 pr
14	67	39.4	300	21	Y77458	Human TNF receptor
15	58	34.1	153	20	Y22222	Human TNFR superfa
16	57	33.5	215	20	W93585	Human hAPO6 protei
17	7	4.1	94	20	Y35734	Amino acid sequenc
18	7	4.1	208	20	Y31236	Human Apo B protei
19	7	4.1	291	8	P70659	Sequence of the N-
20	7	4.1	352	20	Y35164	Amino acid sequenc
21	7	4.1	435	8	P70661	Sequence of fusion
22	7	4.1	447	8	P70660	Sequence of the N-
23	7	4.1	590	19	W62832	Gossypium hirsutum
24	7	4.1	655	21	Y59207	B. halodurans clone
25	7	4.1	809	20	Y09371	Human apolipoprote
26	7	4.1	1035	20	Y49128	Chimeric hmGlur2/h
27	7	4.1	1394	20	Y49129	pmGlur2/Car*Galpa
28	7	4.1	1397	20	Y49134	pmGlur2/Car*Galpa
29	7	4.1	1611	18	W22604	Tylosine synthase
30	7	4.1	1671	11	R06341	Signal peptide and
31	7	4.1	2721	8	P70647	Sequence of N-term
32	6	3.5	7	21	Y60941	Cadherin-5 cell ad
33	6	3.5	8	21	Y60942	Cadherin-5 cell ad
34	6	3.5	9	18	W10275	Antiphospholipid a
35	6	3.5	9	19	W67101	Antibody ACA-6501
36	6	3.5	9	19	W46066	Antiphospholipid a
37	6	3.5	9	21	Y60943	Cadherin-5 cell ad
38	6	3.5	10	21	Y60946	Cadherin-5 cell ad
39	6	3.5	11	21	Y60949	Cadherin-5 cell ad
40	6	3.5	13	18	W30936	Birch pollen aller
41	6	3.5	16	18	W31042	Mugwort pollen all
42	6	3.5	37	20	W88882	Polypeptide fragme
43	6	3.5	40	20	Y41113	TNFR/NGFR cysteine
44	6	3.5	42	20	Y11899	Human 5' EST secre
45	6	3.5	43	20	Y12305	Human 5' EST secre
46	6	3.5	44	11	R07283	Smooth muscle myos
47	6	3.5	50	20	Y27662	Human secreted pro
48	6	3.5	53	19	Y20790	Human neurofilamen
49	6	3.5	57	20	Y13123	Human secreted pro
50	6	3.5	58	19	Y21183	Human bel2 proto-o
51	6	3.5	58	20	W67821	Human secreted pro
52	6	3.5	62	20	Y12367	Human 5' EST secre
53	6	3.5	67	21	Y52254	A. thaliana yg10 p
54	6	3.5	69	20	Y11894	Human 5' EST secre
55	6	3.5	70	21	Y75953	Murine skin cell p
56	6	3.5	92	18	W41642	Rat programmed cel
57	6	3.5	92	18	W41643	Mouse programmed c
58	6	3.5	92	21	Y69865	DP5 protein. Ratt
59	6	3.5	94	20	Y02936	Fragment of human
60	6	3.5	96	20	Y12206	Human 5' EST secre
61	6	3.5	98	12	R14681	FlxX ferredoxin.
62	6	3.5	101	16	R77251	Guinea pig G prote
63	6	3.5	104	20	Y12207	Human 5' EST secre
64	6	3.5	106	19	Y20986	Human g1al fibril
65	6	3.5	118	13	R27755	Overlap of Norwalk
66	6	3.5	118	18	W20129	H. pylori surfac
67	6	3.5	125	21	Y84670	Amino acid sequenc
68	6	3.5	126	17	R91220	Mouse pancreas G-p
69	6	3.5	126	18	W31381	Mouse MIN6-derived
70	6	3.5	126	20	W97223	MIN6-derived.G pro
71	6	3.5	126	20	W95183	Mouse MIN6-derived
72	6	3.5	128	21	Y73489	Human secreted pro
73	6	3.5	133	20	Y27816	Human secreted pro
74	6	3.5	141	21	Y73438	Human secreted pro
75	6	3.5	142	16	R75758	BAV3 ORF3 product.
76	6	3.5	142	18	Y11260	Streptococcus pneu
77	6	3.5	146	20	Y59845	Human normal uteru
78	6	3.5	150	20	Y23765	A human phosphodie
79	6	3.5	151	19	W38678	S. pneumoniae SpO0
80	6	3.5	154	20	Y37855	Amino acid sequenc
81	6	3.5	156	14	R40840	Glioblastoma deriv
82	6	3.5	157	21	Y99443	Human PR01758 (UNQ
83	6	3.5	158	19	Y86114	S. pneumoniae deri
84	6	3.5	180	19	W69664	Human LAGE-1 clone
85	6	3.5	180	21	Y70860	Human LAGE-1 splc

86	6	3.5	183	7	P60624	Sequence B encoded	159	6	3.5	382	19	W76190	Actinoplanes sp. a
87	6	3.5	186	21	W79206	Soluble herpesviru	160	6	3.5	394	18	W25722	Human partial beta
88	6	3.5	191	20	W73377	Human HPDV78 prot	161	6	3.5	398	12	R10572	B. subtilis rib ORF
89	6	3.5	193	21	W79204	Soluble herpesviru	162	6	3.5	398	20	W81552	Bacillus subtilis
90	6	3.5	197	21	W79205	Soluble herpesviru	163	6	3.5	399	20	Y22573	Bacterial general
91	6	3.5	206	17	R91219	Mouse pancreas G-p	164	6	3.5	401	20	Y03143	Heparin sulphate 6
92	6	3.5	206	18	W31380	Mouse MIN6-derived	165	6	3.5	401	20	Y03144	Heparin sulphate 6
93	6	3.5	206	20	W72222	MIN6-derived G pro	166	6	3.5	401	21	Y83903	Mouse HS6ST3 prote
94	6	3.5	206	20	W95182	Mouse MIN6-derived	167	6	3.5	408	20	Y41111	Human TANGO 129 (T
95	6	3.5	210	19	W69663	Human LAGE-1 clone	168	6	3.5	410	19	W68414	Mycobacterium bovi
96	6	3.5	210	21	Y70861	Human LAGE-1 unspl	169	6	3.5	410	20	Y03140	Heparin sulphate 6
97	6	3.5	212	12	R12521	B cell differentia	170	6	3.5	410	21	Y83900	Human HS6ST1 prote
98	6	3.5	219	17	R77660	HPV E6-binding pro	171	6	3.5	411	21	W83358	Mycobacterium tube
99	6	3.5	225	21	Y84798	Amino acid sequenc	172	6	3.5	416	19	W60836	Human LAMP family
100	6	3.5	233	18	W27511	Human hypothalamic	173	6	3.5	416	20	Y29605	Human TSC403 prote
101	6	3.5	235	20	Y73974	Human prostate tum	174	6	3.5	428	21	Y94897	Human protein clon
102	6	3.5	238	20	Y38577	Neisseria meningit	175	6	3.5	430	20	Y41110	Human TANGO 129 (T
103	6	3.5	238	21	Y74979	Neisseria meningit	176	6	3.5	430	21	Y70785	Human tumour necro
104	6	3.5	253	20	Y08025	Mouse protease-rel	177	6	3.5	431	20	Y36903	Protein involved i
105	6	3.5	258	18	W32617	Onion yellow dwarf	178	6	3.5	433	21	Y80093	Permease feci amin
106	6	3.5	275	21	Y81826	Pig lung protease.	179	6	3.5	434	20	Y15215	Amino acid sequenc
107	6	3.5	276	15	R47196	Glutamate racemase	180	6	3.5	437	20	W67722	Human tumour antiq
108	6	3.5	276	18	W20675	H. pylori surface	181	6	3.5	437	21	Y87261	Human signal pepti
109	6	3.5	277	21	Y79207	Membrane-bound her	182	6	3.5	437	21	Y66737	Membrane-bound pro
110	6	3.5	280	19	W72134	HSV-2 strain SB5 C	183	6	3.5	448	21	Y44840	Human orphan cytox
111	6	3.5	283	17	W05809	Human tumour necro	184	6	3.5	456	19	W72032	HSV-2 strain SB5 C
112	6	3.5	283	18	W12659	Human herpes simpl	185	6	3.5	458	20	Y28604	L. cuprina Ecr par
113	6	3.5	283	19	W69238	Herpesvirus entry	186	6	3.5	463	21	Y52216	Arabidopsis thalia
114	6	3.5	283	19	W60045	Human TNF receptor	187	6	3.5	467	21	Y54602	AV37 antigen prote
115	6	3.5	283	20	Y06488	Human tumour-assoc	188	6	3.5	472	18	W23598	Human LYST2 polype
116	6	3.5	283	20	Y05797	Herpes virus entry	189	6	3.5	473	20	Y21210	Human LYST-2 prote
117	6	3.5	283	20	W95031	Tumour necrosis fa	190	6	3.5	473	13	R21977	RuBPCase large sub
118	6	3.5	283	20	W87591	Human tumour necro	191	6	3.5	478	20	Y14051	G. oxydans D-sorbi
119	6	3.5	283	21	Y92518	Human OXRE-15. Ho	192	6	3.5	480	11	R06240	Soluble intercellu
120	6	3.5	284	18	W25165	Mutated tomato mot	193	6	3.5	483	21	Y91433	Human secreted pro
121	6	3.5	287	14	R34030	Fc-alpha-R. Homo	194	6	3.5	487	21	Y75687	Neisseria gonorrh
122	6	3.5	287	15	R59920	Human Fc-alpha-R.	195	6	3.5	487	21	Y75688	Neisseria meningit
123	6	3.5	287	20	Y33189	Human sFc-alphaRI	196	6	3.5	487	21	Y75689	Neisseria meningit
124	6	3.5	293	8	P70405	ORF 2 gene product	197	6	3.5	497	20	W95461	L. helveticus pept
125	6	3.5	293	18	W25166	Spontaneously muta	198	6	3.5	498	15	R55800	Interleukin 14. H
126	6	3.5	293	18	W25167	Spontaneously muta	199	6	3.5	499	20	Y34122	Human potassium ch
127	6	3.5	295	21	Y68463	Mouse testis speci	200	6	3.5	508	18	W14721	Human ICAM-1 (del4
128	6	3.5	296	20	Y19939	B. burgdorferi ant	201	6	3.5	515	15	R48670	Chitinase derivati
129	6	3.5	297	20	W97733	Corn threonine syn	202	6	3.5	518	20	W97737	Soybean threonine
130	6	3.5	308	19	W60946	Streptococcus pneu	203	6	3.5	530	18	W28501	Birch pollen co-fa
131	6	3.5	309	13	R27754	Norwalk virus 8FII	204	6	3.5	531	16	R80110	ICAM-1 CD54. Homo
132	6	3.5	309	17	R98528	Cyclase encoded by	205	6	3.5	532	10	P91357	Inter-cellular adhe
133	6	3.5	311	20	Y41735	Human PRO114 prot	206	6	3.5	532	11	R04165	Inter-cellular adhe
134	6	3.5	311	20	Y39472	DNAX interferon-II	207	6	3.5	532	13	R20809	Inter-cellular adhe
135	6	3.5	311	21	Y44664	Interferon Recepto	208	6	3.5	532	14	R35071	ICAM-1. Homo sapi
136	6	3.5	311	21	Y66676	Membrane-bound pro	209	6	3.5	532	15	R46066	Human ICAM-1. Hom
137	6	3.5	313	13	R27753	Norwalk virus 8FII	210	6	3.5	532	15	R58779	Inter-cellular adh
138	6	3.5	314	20	Y31222	L. vulgareis dFPase	211	6	3.5	532	16	R79457	ICAM-1. Homo sapi
139	6	3.5	317	18	W21949	E6-binding protein	212	6	3.5	532	17	R91437	Human ICAM-1. Hom
140	6	3.5	319	20	Y04749	Mycobacterium spec	213	6	3.5	532	17	R50294	Intracellular adhe
141	6	3.5	323	20	Y19938	B. burgdorferi ant	214	6	3.5	532	18	W27270	Human intracellular
142	6	3.5	340	20	Y42127	Rice cryptophan sy	215	6	3.5	532	18	W14720	Human ICAM-1. Hom
143	6	3.5	346	18	W23088	Pyrodictium sp. es	216	6	3.5	532	18	W09313	Human ICAM-1 (enco
144	6	3.5	369	20	Y16629	Protein encoded b	217	6	3.5	532	19	W80446	Human intracellular
145	6	3.5	369	20	Y07080	Renal cancer assoc	218	6	3.5	532	19	W70871	Intracellular adhe
146	6	3.5	370	17	R91218	Human pituitary G-	219	6	3.5	532	19	W71263	Human intercellula
147	6	3.5	370	18	W31379	Human G protein-co	220	6	3.5	532	19	W46735	Amino acid sequenc
148	6	3.5	370	20	W97221	Human G protein-de	221	6	3.5	532	20	W86193	Human intracellular
149	6	3.5	370	20	W95181	Human G-protein co	222	6	3.5	532	21	Y59499	Human ICAM-1 prote
150	6	3.5	370	21	Y87503	Murine G coupled-p	223	6	3.5	536	19	W81732	M. tuberculosis im
151	6	3.5	370	21	Y87505	Human G coupled-pr	224	6	3.5	536	19	W64365	Mycobacterium tube
152	6	3.5	370	21	Y87507	Rat G coupled-prot	225	6	3.5	536	20	Y39162	M. tuberculosis an
153	6	3.5	374	21	Y75494	Neisseria meningit	226	6	3.5	536	20	Y39019	M. tuberculosis re
154	6	3.5	374	21	Y75728	Neisseria meningit	227	6	3.5	542	20	Y26973	Arabidopsis thalia
155	6	3.5	377	21	Y75492	Neisseria gonorrh	228	6	3.5	543	19	W72196	HSV-2 strain SB5 C
156	6	3.5	377	21	Y75493	Neisseria meningit	229	6	3.5	554	16	R80444	Human soluble epox
157	6	3.5	377	21	Y75727	Neisseria meningit	230	6	3.5	556	21	Y87342	Human signal pepti
158	6	3.5	380	18	W27510	Consensus human hy	231	6	3.5	557	18	W28504	Mugwort grass poll

[illegible]

QY 121 AHAGFCLHSCPPGAGVIAPGESWARGAPRSGRRGQVAGPSLAP 170
 ID Y28449 standard; Protein; 245 AA.
 Db 121 ahagfclhscppgagviapgeswargaprsgrgrrgqvgagpslap 170

RESULT 2
 Y28449
 XX Y28449 standard; Protein; 245 AA.
 AC Y28449;
 XX
 XX 29-SEP-1999 (first entry)
 DT
 DE A human tumour necrosis factor-R2-like proteins (TR2P)-1.
 XX
 KW Human tumour necrosis factor-R2-like protein; TR2P; achondroplasia;
 KW osteoporosis; developmental disorder; Cushing's syndrome;
 KW muscular dystrophy; epilepsy; hereditary neuropathy;
 KW Charcot-Marie-Tooth disease; neurofibromatosis; hypothyroidism;
 KW hydrocephalus; seizure disorder; cerebellar palsy; spinal bifida;
 KW congenital glaucoma; cataract; sensorineural hearing loss;
 KW reproductive disorder; infertility; ovulatory defect; endometriosis;
 KW autoimmune disorder; ectopic pregnancy; teratogenesis; spermatogenesis;
 KW immunological disorder; AIDS; Addison's disease; allergy; bronchitis;
 KW atherosclerosis; diabetes mellitus; Chron's disease; lupus;
 KW irritable bowel syndrome; multiple sclerosis; infection;
 KW neoplastic disorder; adenocarcinoma; leukaemia; lymphoma; melanoma;
 KW myeloma; sarcoma.
 XX
 OS Homo sapiens.
 XX
 PN WO9931128-A2.
 XX
 PD 24-JUN-1999.
 XX
 PF 02-DEC-1998; 98WO-US25649.
 XX
 PR 16-DEC-1997; 97US-0991945.
 XX
 XX (INCY-) INCYTE PHARM INC.
 PA
 PI Au-Young J, Bandman O, Hillman JL, Kaser MR, Tang YT;
 XX
 DR WPI; 1999-457916/38.
 DR N-PSDB; X89503.
 XX
 PT New tumour necrosis factor-R2-like protein - useful in the treatment
 of osteogenesis, developmental, reproductive, immunological and
 neoplastic disorders
 PT
 XX
 PS Claim 1; Fig 1A-C; 81pp; English.
 XX
 CC The present sequence represents a human tumour necrosis factor-R2-like
 protein (TR2P). The protein is used to treat and prevent osteogenesis,
 CC developmental, reproductive, immunological and neoplastic disorders, and
 CC also to diagnose disorders associated with TR2 protein expression. Such
 CC disorders include osteogenesis disorders such as achondroplasia and
 CC osteoporosis, developmental disorders such as Cushing's syndrome,
 CC muscular dystrophy, and epilepsy, hereditary neuropathies such as
 CC Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism,
 CC hydrocephalus, seizure disorders such as cerebral palsy and spinal
 CC bifida, congenital glaucoma, cataract, or sensorineural hearing loss,
 CC reproductive disorders such as infertility, ovulatory defects and
 CC endometriosis, autoimmune disorders, ectopic pregnancy and teratogenesis,
 CC disruption of spermatogenesis, immunological disorders such as AIDS,
 CC Addison's disease, allergies, bronchitis, atherosclerosis, diabetes
 CC mellitus, Chron's disease, lupus and irritable bowel syndrome, multiple
 CC sclerosis, viral, fungal, helminthic, paracitic and protozoal infections,
 CC and neoplastic disorders including adenocarcinoma, leukaemia, lymphoma,
 CC melanoma, myeloma, sarcoma, and teratocarcinoma.
 XX
 SQ Sequence 245 AA;

Query Match 83.5%; Score 142; DB 20; Length 245;
 Best Local Similarity 100.0%; Pred. No. 1.8e-131;
 Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALEGPGLSLCLVLALPALLPFAVRGVAETPYVNRDAETGRLVCAQCQPPGTFVQR 60
 ID W66102 standard; Protein; 300 AA.
 Db 1 mralepglslclvlalpallpavrgvaetptyvnrdaetgerlvcaqcpptgfvr 60

QY 61 PCRRDSPPTTCGCPPRHYTFQFNWYLERCRYCNVLCGEREEERACHATHNRACRTGFF 120
 ID W66102 standard; Protein; 300 AA.
 Db 61 pccrrdsppttcgcpprhytfqfnwylercrycnvlgereeeerachathnracrtgff 120

QY 121 AHAGFCLHSCPPGAGVIAPG 142
 ID W66102 standard; Protein; 300 AA.
 Db 121 ahagfclhscppgagviapg 142

RESULT 3
 W66102
 ID W66102 standard; Protein; 300 AA.
 XX
 AC W66102;
 XX
 DT 02-DEC-1998 (first entry)
 XX
 DE Amino acid sequence of tumour necrosis related receptor (TR4).
 XX
 KW Human; tumour necrosis related receptor; TR4; agonist; antagonist;
 KW inhibition; chronic; acute; inflammation; arthritis; septicaemia;
 KW autoimmune disease; transplant rejection; stroke; cancer;
 KW Alzheimer's disease.
 XX
 OS Homo sapiens.
 XX
 PN EP861850-A1.
 XX
 PD 02-SEP-1998.
 XX
 PF 20-JAN-1998; 98EP-0300382.
 XX
 PR 04-FEB-1997; 97US-0794796.
 XX
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 PA
 PI Emery J, Tan KB, Truneh A, Young PR;
 XX
 DR WPI; 1998-508248/44.
 DR N-PSDB; V07654.
 XX
 PT New DNA encoding tumour necrosis related receptor - used to treat
 PT and prevent e.g. inflammation, arthritis, septicaemia, autoimmune
 PT diseases, transplant rejection, infection, stroke, ischaemia, ARDS,
 PT restenosis, AIDS, bone disorders and cancer
 PS
 XX Claim 1; Fig 1; 21pp; English.
 XX
 CC This is the amino acid sequence of the human tumour necrosis related
 CC receptor (TR4), used in the method of the invention. The TR4 protein
 CC or its agonist can be used to treat a subject in need of enhanced
 CC TR4 polypeptide activity. The antagonist is used to inhibit TR4
 CC polypeptide activity. The active agents can be used for the
 CC treatment and prevention of diseases such as chronic and acute
 CC inflammation, arthritis, septicaemia, autoimmune diseases, transplant
 CC rejection, stroke, cancer, Alzheimer's disease.
 XX
 SQ Sequence 300 AA;

Query Match 83.5%; Score 142; DB 19; Length 300;
 Best Local Similarity 100.0%; Pred. No. 2.1e-131;
 Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRALEGPGLSLCLVLALPALLPVPVARGVAETPTYPWRDAETGERLVCAOCPGCTFVQR 60
 |||||
 Db 1 mralepgpalslclvlalpalppavrgvaetpypwrdaetgerlvcaocppgtfvqr 60
 |||||
 Qy 61 PCRRDSPPTCGCPPRHHYTFQWNYLERCRYCNVLCGEREEERACHATHNRACRRTGFF 120
 |||||
 Db 61 pcrddspptcgcprhhytqfwnylercrycnvlgereeeearachathnrcrctgff 120
 |||||
 Qy 121 AHAGFCLHASCPCPGAGVIAPG 142
 |||||
 Db 121 ahagfclenascpvgagviapg 142
 |||||

RESULT 4
 W63622
 ID W63622 standard; Protein; 300 AA.
 XX
 AC W63622;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Human tumour necrosis factor receptor-6 alpha protein.
 XX
 KW Human tumour necrosis factor receptor-6 alpha; TNFR-6 alpha; TNFR-6 beta;
 KW endothelial cells; keratinocytes; normal prostate; apoptosis;
 KW prostate tumour tissue.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..30
 FT Protein 31..300
 FT /note= "TNFR-6 alpha"
 FT Region 31..282
 FT /note= "Soluble extracellular domain"
 XX
 XX WO9830694-A2.
 PN
 PD 16-JUL-1998.
 XX
 PF 13-JAN-1998; 98WO-US00153.
 XX
 PR 14-JAN-1997; 97US-0035496.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ebner R, Feng P, Gentz RL, Ni J, Ruben SM, Yu G;
 DR WPI; 1998-399142/34.
 DR N-PSDB; V39085.
 XX
 PT Human tumour necrosis factor receptors 6-alpha and 6-beta - used in
 PT the diagnosis of immune system-related disorder(s)
 XX
 PS Claim 20; Fig 1; 91pp; English.
 XX
 CC The present sequence represents the human tumour necrosis factor
 CC receptor-6 alpha (TNFR-6 alpha) protein. The invention also provides
 CC for the TNFR-6 beta protein (W63623). TNFR-6 alpha and TNFR-6 beta are
 CC members of the tumour necrosis factor receptor (TNFR) family. TNFRs
 CC are expressed in endothelial cells, keratinocytes, normal prostate and
 CC prostate tumour tissue. For a number of disorders of these cells,
 CC particularly of the immune system, substantially altered (whether
 CC increased or decreased) levels of TNFR-6 alpha and/or TNFR-6 beta gene
 CC expression can be detected, therefore the TNFR-6 alpha and TNFR-6 beta
 CC polypeptides, nucleic acids and antibodies are claimed to be useful in
 CC the diagnosis of such disorders. Mutations of the TNFR-6 alpha and
 CC TNFR-6 beta genes can also be detected. The TNFR polypeptides are
 CC also claimed to be useful for identifying ligands which may be useful
 CC in the treatment of apoptosis related disorders.
 XX
 SQ Sequence 300 AA;

Query Match 83.5%; Score 142; DB 19; Length 300;
 Best Local Similarity 100.0%; Pred. No. 2.1e-131;
 Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRALEGPGLSLCLVLALPALLPVPVARGVAETPTYPWRDAETGERLVCAOCPGCTFVQR 60
 |||||
 Db 1 mralepgpalslclvlalpalppavrgvaetpypwrdaetgerlvcaocppgtfvqr 60
 |||||
 Qy 61 PCRRDSPPTCGCPPRHHYTFQWNYLERCRYCNVLCGEREEERACHATHNRACRRTGFF 120
 |||||
 Db 61 pcrddspptcgcprhhytqfwnylercrycnvlgereeeearachathnrcrctgff 120
 |||||
 Qy 121 AHAGFCLHASCPCPGAGVIAPG 142
 |||||
 Db 121 ahagfclenascpvgagviapg 142
 |||||

RESULT 5
 Y03099
 ID Y03099 standard; Protein; 300 AA.
 XX
 AC Y03099;
 XX
 DT 09-DEC-1999 (first entry)
 XX
 DE Human lung TNF-receptor protein.
 XX
 KW Tumour necrosis factor; TNF; TNF receptor; human; lung; gene therapy;
 KW detection; immunoassay; diagnosis; disease; immune system; tumour;
 KW osteogenic system; cardiovascular system; central nervous system; asthma;
 KW peripheral nervous systems; transplant incompatibility; antitumor;
 KW rheumatoid arthritis; antiasthmatic; antiarthritic.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 134..1036
 FT /*tag= a
 FT /product= "TNF-receptor"
 XX
 PN DE19809978-A1.
 XX
 PD 16-SEP-1999.
 XX
 PF 09-MAR-1998; 98DE-1009978.
 XX
 PR 09-MAR-1998; 98DE-1009978.
 XX
 PA (BADI) BASF AG.
 XX
 PI Kroeger B;
 DR WPI; 1999-519473/44.
 DR N-PSDB; Z09998.
 XX
 PT New soluble member of tumor necrosis factor receptor family, useful for
 PT identification specific modulators and for treating disease e.g. tumors
 XX
 PS Claim 1; Page 8-9; 10pp; German.
 XX
 CC This invention describes a novel tumour necrosis factor (TNF) receptor
 CC (I) isolated from human lung tissue. (I) is used: (i) to raise specific
 CC antibodies (Ab); (ii) to screen for specific (ant)agonists or ligands
 CC (A), potential therapeutic agents; and (iii) therapeutically (optionally
 CC expressed from a gene therapy vector) in conditions associated with a
 CC deficit of (I). Ab are used: (a) for qualitative or quantitative
 CC detection of (I) in standard immunoassays (for diagnosis of disease, or
 CC susceptibility, or for monitoring); and (b) as therapeutic inhibitors in
 CC cases where (I) is overexpressed. Nucleic acid (II) that encodes (I) is
 CC used: (A) for recombinant production of (I); (B) also its oligonucleotide

CC fragments, in standard hybridization and/or amplification assays; (C) as
 CC source of antisense molecules or ribozymes; and (D) to produce transgenic
 CC animals (for studying (patho)physiology of (I)). Diseases possibly
 CC associated with under- or over-expression of (I) are those of the immune,
 CC osteogenic, cardiovascular and central or peripheral nervous systems,
 CC tumors, transplant incompatibility, asthma and rheumatoid arthritis. The
 CC products of the invention have antitumor, antiasthmatic and
 CC antiarthritic activity. This sequence represents the TNF-receptor of the
 CC invention.
 XX
 SQ Sequence 300 AA;

Query Match 83.5%; Score 142; DB 20; Length 300;
 Best Local Similarity 100.0%; Pred. No. 2.1e-131;
 Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALEGGLSLCLVLALPALLPVPVAVGVAETPTYPWRDAETGERLVCAQCPGTFVQR 60

DB 1 mralepgllslclvlalpallpvpavrgvaeetptypwrdaetgerlvcaqcpptfvqr 60

QY 61 PCRRDSPTTCGCPGPPRHYYTFWNYLRCRYCNVLCGEREEERACHATHNRACRRTGFF 120

DB 61 pcrdspttcgpcpprhytqfwnylrcrycnvlgereeeearachathnrcrctgff 120

QY 121 AHAGFCLHASCPGAGVIAPG 142

DB 121 ahagfclhascpaggaviapg 142

RESULT 6

Y42182 ID Y42182 standard; Protein; 300 AA.

AC Y42182;

XX DT 17-DEC-1999 (first entry)

XX DE Human FLINT #1 protein sequence.

XX KW Human; FLINT; mFLINT; OPC3; tumour necrosis factor receptor; FasL;
 KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
 KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
 KW reperfusion-associated injury; aplastic anaemia; differentiation;
 KW growth; myelodysplastic syndrome; pancytopenic condition;
 KW myocardial ischaemia.

XX OS Homo sapiens.

XX PN W09950413-A2.

XX PD 07-OCT-1999.

XX PF 30-MAR-1999; 99WO-US06797.

XX PR 30-MAR-1998; 98US-0079856.

XX PR 20-MAY-1998; 98US-0086074.

XX PR 09-SEP-1998; 98US-0099643.

XX PR 17-DEC-1998; 98US-0112577.

XX PR 18-DEC-1998; 98US-0112703.

XX PR 18-DEC-1998; 98US-0112933.

XX PR 22-DEC-1998; 98US-0113407.

XX PA (ELIL) LILLY & CO ELI.

XX PI Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;

PI Hui KY, Khaitonenkov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
 PI Song HY, Wang J, Wu X, Zuckerman SH;

XX WPI; 1999-591319/50.
 XX DR N-PSDB; 225375.

XX PT Use of mature FLINT for treating acute liver failure, inflammation,

PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
 PT and proinflammatory activity;
 XX
 PS Claim 30; Fig 1; 99pp; English.

XX The present invention describes therapeutic applications of mature FLINT
 CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
 CC (mFLINT), which is a member of the tumour necrosis factor receptor
 CC superfamily, is used for treating acute liver failure, inflammation of
 CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
 CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
 CC injury or disorder such as hypercoagulation (including use with
 CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
 CC or disorder, Type I diabetes, cancer, cell damage or damage to an
 CC innocent bystander tissue that is induced by a chemotherapeutic agent or
 CC therapeutic irradiation, treating haematopoietic progenitor cells that
 CC have been exposed to therapeutic radiation or chemotherapy, aplastic
 CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
 CC also used for promoting the growth or differentiation of a haematopoietic
 CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
 CC resulting from abnormal myocardial ischaemia. The present sequence
 CC represents human FLINT.

XX Sequence 300 AA;

Query Match 83.5%; Score 142; DB 20; Length 300;

Best Local Similarity 100.0%; Pred. No. 2.1e-131;
 Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALEGGLSLCLVLALPALLPVPVAVGVAETPTYPWRDAETGERLVCAQCPGTFVQR 60

DB 1 mralepgllslclvlalpallpvpavrgvaeetptypwrdaetgerlvcaqcpptfvqr 60

QY 61 PCRRDSPTTCGCPGPPRHYYTFWNYLRCRYCNVLCGEREEERACHATHNRACRRTGFF 120

DB 61 pcrdspttcgpcpprhytqfwnylrcrycnvlgereeeearachathnrcrctgff 120

QY 121 AHAGFCLHASCPGAGVIAPG 142

DB 121 ahagfclhascpaggaviapg 142

RESULT 7

Y17479 ID Y17479 standard; Protein; 300 AA.

XX AC Y17479;

XX DT 02-AUG-1999 (first entry)

XX DE Mammalian tumour necrosis factor receptor OPG-2.

XX KW Tumour necrosis factor receptor; TNF receptor; OPG-2; Paget's disease;
 KW osteopenic disorder; osteoclast activity; primary osteoporosis;
 KW hyperglycaemia; osteolytic metastasis; immune response; cancer.

XX OS Mammalia.

XX PN W09926977-A1.

XX PD 03-JUN-1999.

XX PF 24-NOV-1998; 98WO-US25065.

XX PR 17-FEB-1998; 98US-0074896.

XX PR 24-NOV-1997; 97US-0066446.

XX PA (BIOJ) BIOGEN INC.

XX PI Tschopp J;

XX WPI; 1999-347693/29.

DR N-PSDB; X76052.
XX
PT New tumour necrosis factor family receptor OPG-2
XX
PS Claim 1; Page 18; 22pp; English.
XX
CC The present sequence represents a mammalian tumour necrosis factor
CC receptor, designated OPG-2. OPG-2, is a member of the tumour necrosis
CC factor receptor family, and can be used: (i) to raise specific
CC antibodies (Ab), (ii) to treat osteopenic disorders associated with
CC excessive osteoclast activity, e.g. primary osteoporosis, Paget's
CC disease, hyperglycaemia of malignancy, or osteolytic metastases; (iii)
CC for affinity purification of cognate ligands, and (iv) to screen for
CC ligands (antagonists or agonists). Ab, or other OPG-2 blocking agents
CC such as soluble forms of the protein, are used to prevent, or reduce
CC severity of, an immune response, and for treating cancer. They can also
CC be used in diagnostic assays. The nucleic acid sequence encoding OPG-2
CC can be used as a probe to isolate related sequences from other species.
XX
SQ Sequence 300 AA;

Query Match 83.5%; Score 142; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.le-131;
Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRALEGPSLSLCLVLPALLPVPVAVRGVAETPTYPWRDAETGERLVCAOCPGCTFVQR 60
Db 1 mraleppgslslclvlalpallpvpvrgvvaetpypwrdaetgerlvcaqcpptfvqr 60

Qy 61 PCRDSPTTCGCPPRHYTQFWNYLRCRYCNVLCGEREEERACHATHNRCRCRTGFF 120
Db 61 pcrdpspttcgpcprhytqfwnylercrycnvlgereeeearachathnrcrcrtgff 120

Qy 121 AHAGFCLHASCPCGAGVIAPG 142
Db 121 ahagfclhascpbgagviapg 142

RESULT 8
Y06817
ID Y06817 standard; Protein; 300 AA.
XX
AC Y06817;
XX
XX 24-JUN-1999 (first entry)
XX
XX Human Dcr3 polypeptide.
XX
KW Dcr3 polypeptide; tumour necrosis factor receptor; TNFR; Fas ligand;
KW apoptosis; T cell mediated immune response; allergy; asthma; cancer;
KW rheumatoid arthritis; Crohn's disease; guest vs. host disease; human;
KW gene therapy.
XX
OS Homo sapiens.
XX
FN WO9914330-A1.
XX
PD 25-MAR-1999.
XX
XX 18-SEP-1998; 98WO-US19661.
XX
XX 30-JUL-1998; 98US-0094640.
XX
PR 18-SEP-1997; 97US-0059288.
XX
XX (GETH) GENENTECH INC.
XX
XX Ashkenazi AJ, Botstein D, Dodge KH, Goddard A, Gurney AL;
PI Kim KJ, Lawrence DA, Pitti R, Roy MA, Tumas DB;
PI Wood WI;
XX
XX WPI; 1999-244032/20.
DR
DR N-PSDB; X32744.

XX Dcr3 polypeptide related to tumor necrosis factor receptor
PT
XX Claim 5; Fig 1; 8pp; English.
PS
XX
CC This represents a human Dcr3 polypeptide, a homologue of tumour necrosis
CC factor receptor (TNFR) polypeptide. Host cells containing a vector
CC comprising the Dcr3 nucleic acid can be used for the recombinant
CC expression of the protein. Dcr3 binds to Fas ligand, so it (or its
CC chimeras) are useful for modulating apoptosis in mammalian cells, also
CC other Fas-ligand induced activities, particularly to inhibit T cell
CC mediated immune responses, e.g. in treatment of allergy, asthma,
CC rheumatoid arthritis, Crohn's disease, guest vs. host disease etc. Dcr3
CC may also be used to identify specific binding proteins, potential
CC inhibitors. Antibodies against Dcr3 are used to treat cancer,
CC specifically of the lung and colon, also in diagnosis and for affinity
CC purification of the protein. Detecting mutations in the gene for Dcr3 is
CC also used to diagnose cancer, or predisposition to it. Dcr3 nucleic acid
CC is useful as hybridization probe to detect genomic or related sequences;
CC for chromosome and gene mapping; as source of antisense sequences; for
CC expression of recombinant Dcr3 and to generate transgenic animals (for
CC development and screening of therapeutic agents), also for in vivo or
CC ex vivo gene therapy.
XX
SQ Sequence 300 AA;

Query Match 83.5%; Score 142; DB 20; Length 300;
Best Local Similarity 100.0%; Pred. No. 2.le-131;
Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRALEGPSLSLCLVLPALLPVPVAVRGVAETPTYPWRDAETGERLVCAOCPGCTFVQR 60
Db 1 mraleppgslslclvlalpallpvpvrgvvaetpypwrdaetgerlvcaqcpptfvqr 60

Qy 61 PCRDSPTTCGCPPRHYTQFWNYLRCRYCNVLCGEREEERACHATHNRCRCRTGFF 120
Db 61 pcrdpspttcgpcprhytqfwnylercrycnvlgereeeearachathnrcrcrtgff 120

Qy 121 AHAGFCLHASCPCGAGVIAPG 142
Db 121 ahagfclhascpbgagviapg 142

RESULT 9
W97749
ID W97749 standard; Protein; 300 AA.
XX
AC W97749;
XX
XX 21-MAY-1999 (first entry)
XX
XX Human tumour necrosis factor receptor ZTNFR-5.
XX
XX ZTNFR-5; tumour necrosis factor receptor; TNFR; human;
KW cell maturation; bone cell regulation.
XX
XX Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT Peptide 1..23 /note= "signal peptide"
FT Protein 24..300 /note= "mature protein"
FT Domain 24..194 /note= "extracellular domain"
FT Region 49..71 /note= "cysteine-rich pseudo-repeat 1"
FT Region 72..113 /note= "cysteine-rich pseudo-repeat 1"
FT Region 114..151 /note= "cysteine-rich pseudo-repeat 1"
FT Region 152..194 /note= "cysteine-rich pseudo-repeat 1"

RESULT 13
 Y42185
 ID Y42185 standard; Protein; 273 AA.
 XX
 AC Y42185;
 XX
 DT 17-DEC-1999 (first entry)
 XX
 DE Human mFLINT #2 protein sequence.
 XX
 KW Human; FLINT; mFLINT; OPG3; tumour necrosis factor receptor; FasL;
 KW apoptosis; inflammation; cancer; diabetes; acute liver failure;
 KW sepsis; hepatitis; ischaemia-associated injury; hypercoagulation;
 KW reperfusion-associated injury; aplastic anaemia; differentiation;
 KW growth; myelodysplastic syndrome; pancytopenic condition;
 KW myocardial ischaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO9950413-A2.
 XX
 PD 07-OCT-1999.
 XX
 PF 30-MAR-1999; 99WO-US06797.
 XX
 PR 30-MAR-1998; 98US-0079856.
 PR 20-MAY-1998; 98US-0086074.
 PR 09-SEP-1998; 98US-0099643.
 PR 17-DEC-1998; 98US-0112577.
 PR 18-DEC-1998; 98US-0112703.
 PR 18-DEC-1998; 98US-0112933.
 PR 22-DEC-1998; 98US-0113407.
 XX
 PA (ELIL) LILLY & CO ELI.
 XX
 PI Bumol TF, Dou S, Glasebrook AL, Gould KE, Hale JE, Heuer JG;
 PI Hui KY, Kharitonov A, Mizrahi J, Na S, Noblitt TW, Reidy CA;
 PI Song HY, Wang J, Wu X, Zuckerman SH;
 XX
 WPI: 1999-591319/50.
 DR N-PSDB; Z25378.
 XX
 PT Use of mature FLINT for treating acute liver failure, inflammation,
 PT cancer, and diabetes - by prevention of FasL-Fas mediated apoptotic
 PT and proinflammatory activity
 XX
 PS Example 2; Fig 4; 99pp; English.
 XX
 CC The present invention describes therapeutic applications of mature FLINT
 CC (mFLINT) for use in the treatment of acute liver failure. Mature FLINT
 CC (mFLINT), which is a member of the tumour necrosis factor receptor
 CC superfamily, is used for treating acute liver failure, inflammation of
 CC the liver, abnormal hepatocyte apoptosis, sepsis, a disorder associated
 CC with inflammation, hepatitis, abnormal apoptosis, an ischaemia-associated
 CC injury or disorder such as hypercoagulation (including use with
 CC thrombolytic or anti-thrombolytic agents), reperfusion-associated injury
 CC or disorder. Type I diabetes, cancer, cell damage or damage to an
 CC innocent bystander tissue that is induced by a chemotherapeutic agent or
 CC therapeutic irradiation, treating haematopoietic progenitor cells that
 CC have been exposed to therapeutic radiation or chemotherapy, aplastic
 CC anaemia, myelodysplastic syndrome or a pancytopenic condition. mFLINT is
 CC also used for promoting the growth or differentiation of a haematopoietic
 CC progenitor cell or CD34+ cell and preventing damage to a cardiac myocyte
 CC resulting from abnormal myocardial ischaemia. The present sequence
 CC represents human mFLINT.
 XX
 SQ Sequence 273 AA;
 XX
 Query Match 51.2%; Score 87; DB 20; Length 273;
 Best Local Similarity 100.0%; Pred. No. 1.5e-77;
 Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 VAETPTYPWRDAETGERLVCAQCPCPGTFVQRCRRDSTTCGCPGPRHYTQFWNYLERCR 89
 Db 1 vaetptypwrdaetgerlvcaqcpcpgtfvqrcrrdspsttcgpcpprhytqfwnylercr 60
 OY 90 YCNVLGGEREEERACHATHNRACRCR 116
 Db 61 ycnvlgerereearachathnrcrcr 87
 RESULT 14
 Y77458
 ID Y77458 standard; Protein; 300 AA.
 XX
 AC Y77458;
 XX
 DT 05-JUN-2000 (first entry)
 XX
 DE Human TNF receptor-like protein, HDTEA84.
 XX
 KW TNF receptor family; tumour necrosis factor; HDTEA84; HSLJD37R;
 KW Rank-like protein; RANKL; immune disorder; inflammation; allergy;
 KW immunosuppressant; antiarthritic; antirheumatoid; antiinflammatory;
 KW dermatological; antithyroid.
 XX
 OS Homo sapiens.
 XX
 PN WO200001817-A2.
 XX
 PD 13-JAN-2000.
 XX
 PF 06-JUL-1999; 99WO-US12366.
 XX
 PR 06-JUL-1998; 98US-0110938.
 PR 13-JUL-1998; 98US-0114466.
 PR 23-JUL-1998; 98US-0093897.
 PR 12-AUG-1998; 98US-0132968.
 PR 18-AUG-1998; 98US-0136214.
 PR 11-SEP-1998; 98US-0099999.
 XX
 PA (SCHE) SCHERING CORP.
 XX
 PI Bates EEM, Lebecque SJE, Murphy EE, Mattison JD, Gorman DM;
 PI Hedrick JA, Wang L, Zlotnik A, Murgolo NJ, Greene JR, Johnston JA;
 PI Bazan JF, Mahony D, Lees EM;
 XX
 WPI: 2000-171015/15.
 DR N-PSDB; Z92404.
 XX
 PT New isolated mammalian genes, used to develop products for treating
 PT e.g. immune, inflammatory or allergic abnormalities, cancers or
 PT degenerative conditions
 XX
 PS Claim 24; Page 157; 218pp; English.
 XX
 CC The invention relates to a number of primate and/or rodent proteins, and
 CC the genes which encode them. The invention encompasses human dendritic
 CC cell prostaglandin transporter (DC-PGT); the TNF (tumour necrosis
 CC factor) receptor family-related proteins HDTEA84, HSLJD37R and RANKL;
 CC human CC chemokine HCC5; human deubiquitinating proteins Dub1 and Dub
 CC 12; human MD-1 and human and murine MD-2 proteins, which exhibit the
 CC properties of ligands for proteins comprising a leucine-rich motif (LRR);
 CC human cyclin E2; cDNAs encoding these proteins; and antibodies against
 CC these proteins. The proteins can be used for modulating the physiology or
 CC development of a cell. They can be used to mediate uptake of substrates
 CC (e.g., prostaglandin-like molecules), to modulate or mediate cellular
 CC interactions (e.g., induce or prevent trafficking, proliferation, or
 CC differentiation of cells), or are intracellular proteins which are
 CC important in various cellular processes such as the deubiquitination of
 CC proteins or cell cycle regulation. The products can be used for
 CC treating medical conditions such as immune, inflammatory or allergic
 CC disorders, or abnormal cellular proliferation, for example, cancers or
 CC degenerative conditions. They can be used to modulate immune responses in

CC disease states e.g., autoimmune disorders, including rheumatoid
 CC arthritis, systemic lupus erythematosus, Hashimoto's autoimmune
 CC thyroiditis, as well as acute and chronic inflammatory responses in which
 CC T cell activation, expansion, and/or immunological T cell memory play an
 CC important role. Sequences Y77458-Y77461 and Y77465-Y77468 represent
 CC TNF receptor family-related proteins. Y77458 is the human protein
 CC HDTA84, Y77459-Y77461 are human HSLJD37R proteins, Y77465 is
 CC murine Rank-like protein RANKL, and Y77466-Y77468 are human RANKL
 CC proteins.
 XX
 SQ Sequence 300 AA;

Query Match 39.4%; Score 67; DB 21; Length 300;
 Best Local Similarity 100.0%; Pred. No. 6.7e-58;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRALEGPGLSLLCLVLPALLPVPVARGVAETPTYPWRDAETGERLVCAQCPTGVQR 60
 Db 1 mralegpglsllclvlpallpvpavrgvaetptypwrdaetgerlvcaqcpptfvqr 60

QY 61 PCRDPSP 67

Db 61 pcrdpdp 67

RESULT 15

Y22222
 ID Y22222 standard; Protein; 153 AA.

XX
 AC Y22222;

XX
 DT 16-SEP-1999 (first entry)

XX
 DE Human TNFR superfamily soluble receptor protein sequence.

XX TNFL1; human; TNFR superfamily; tumour necrosis factor ligand; TNF;
 KW tumour necrosis factor receptor; TNFR superfamily; cell proliferation;
 KW cell differentiation; cytokine production; immunoglobulin; hyperplasia;
 KW apoptosis inducer; activated T cell; autoimmune disease; inhibitor;
 KW myasthenia gravis; insulin-dependent diabetes mellitus; endotoxin shock;
 KW rheumatoid arthritis; multiple sclerosis; systemic lupus erythematosus;
 KW tumour; proliferative disorder; neoplasia; dysplasia; immunocompetence;
 KW lymphoid organogenesis; bacterial resistance; contact hypersensitivity;
 KW delayed type sensitivity; therapy.

XX Homo sapiens.

XX WO9933980-A2.

XX 08-JUL-1999.

XX 22-DEC-1998; 98WO-US27474.

XX 16-DEC-1998; 98US-0212270.

XX 30-DEC-1997; 97US-0068959.

XX (CHIR) CHIRON CORP.

XX Kassam A, Lamson G, Pot D, Tribouley C;

XX WPI; 1999-405508/34.

XX N-PSDB; X84621.

XX New tumour necrosis factor ligands, useful for induction of cell
 PT death and/or proliferation of cells

XX Claim 1; Page 61; 69pp; English.

XX This sequence represents a tumour necrosis factor receptor (TNFR)
 CC superfamily soluble protein of the invention. The invention also relates
 CC to tumour necrosis factor (TNF) ligand (TNFL) family proteins. The TNFL
 CC proteins play regulatory roles in cell proliferation and/or

CC differentiation, e.g. they can induce production of cytokines,
 CC immunoglobulins, etc. A variety of diseases can be treated by modulating
 CC the activity of TNFL proteins, e.g. they can induce apoptosis of
 CC activated T cells but rescue resting T cell from apoptosis. TNFL
 CC polypeptides can therefore be used to treat autoimmune diseases, such as
 CC myasthenia gravis, insulin-dependent diabetes mellitus, rheumatoid
 CC arthritis, multiple sclerosis, and systemic lupus erythematosus. TNFL
 CC proteins also have tumour stimulating properties, so tumours can be
 CC treated by inhibiting the expression or activity of TNFL. Other
 CC proliferative disorders, such as neoplasias, dysplasias, and hyperplasia
 CC can also be treated using TNFL inhibitors. The TNFL polypeptides and
 CC polynucleotides can also be used to enhance or decrease TNF activity,
 CC thus providing therapeutic benefits such as induction of cell death,
 CC lymphoid organogenesis, or host bacterial resistance, and inhibition of
 CC endotoxin shock, contact hypersensitivity, delayed type sensitivity or
 CC immunocompetence of a transplant recipient. TNF and its receptors play a
 CC major role in host defence and immunosurveillance. As such, there is a
 CC need to identify new members of TNFR families. This invention provides
 CC this need.
 XX
 SQ Sequence 153 AA;

Query Match 34.1%; Score 58; DB 20; Length 153;

Best Local Similarity 100.0%; Pred. No. 2.5e-49;

Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 LERCRCYNLCGEREEERACHATHNRACRCRTGFFAHAGFCLEHASCPPGAGVIAPG 142
 Db 1 lercrcynlcgerereearachathnracrcrtgffahagfclehascpaggviapg 58

RESULT 16

W93585

ID W93585 standard; Protein; 215 AA.

XX W93585;

XX DT 18-JUN-1999 (first entry)

XX DE Human hAPO6 protein.

XX Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
 KW developmental abnormality; gestational abnormality; prostate cancer;
 KW APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
 KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
 KW apoptosis; human.

XX Homo sapiens.

XX WO9911791-A2.

XX 11-MAR-1999.

XX 04-SEP-1998; 98WO-US18393.

XX 05-SEP-1997; 97US-0924634.

XX (UNIW) UNIV WASHINGTON.

XX Chaudhary PM;

XX WPI; 1999-205191/17.

XX N-PSDB; X23419.

XX New Tumour Necrosis Factor family receptor polypeptides and ligands -
 PT useful for diagnosis and treatment of prostate cancer and
 PT developmental or gestational abnormalities

XX Claim 29; Fig 9; 156pp; English.

XX This invention describes isolated Tumour Necrosis Factor (TNF) family
 CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active

CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
 CC their active fragments. APO4 is useful for diagnosing prostate cancer
 CC by determining levels of APO4 in an individual. Prostate cancer can also
 CC be treated using APO4 selective binding agents linked to a therapeutic
 CC moiety. APO4 polypeptides are also useful for identifying selective
 CC binding agents, useful in diagnosis/treatment of disease by binding of
 CC agents to the polypeptide/active fragment which is extracellular, or
 CC expressed on the cell surface. The binding is preferably performed in
 CC vivo. APO4 polypeptides/ active fragments are also useful for screening
 CC for agonists and antagonists by binding and observing the change in APO4
 CC activity. Effective pharmacological agents useful in diagnosis or
 CC treatment of disease are also identified using APO4 polypeptides/active
 CC fragments and APO4 signal transducer molecules that specifically interact
 CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4
 CC activity. The method is performed in vivo or in vitro. APO polypeptides
 CC are all useful as immunogens for preparing antibodies. APO4 is also
 CC useful for diagnosis/treatment of developmental or gestational
 CC abnormalities. APO8 was transfected to human breast carcinoma cell line
 CC MCF-7, and induced apoptosis.
 XX
 SQ Sequence 215 AA;

Query Match 33.5%; Score 57; DB 20; Length 215;
 Best Local Similarity 100.0%; Pred. No. 3.2e-48;
 Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 86 ERCRYNVLGGREERACHATHNRACRCRTGFFAHAGFCLEHASCPPGAGVIAPG 142
 DB 1 ERCRYNVLGGREERACHATHNRACRCRTGFFAHAGFCLEHASCPPGAGVIAPG 57

RESULT 17
 Y35734
 ID Y35734 standard; Protein; 94 AA.
 XX
 AC Y35734;
 XX
 DT 13-SEP-1999 (first entry)
 XX
 DE Amino acid sequence of a Chlamydia pneumoniae protein.
 XX
 KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
 KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
 KW vaccine; neutralising epitope.
 XX
 OS Chlamydia pneumoniae.
 XX
 PN W09927105-A2.
 XX
 PD 03-JUN-1999.
 XX
 PF 20-NOV-1998; 98WO-1B01890.
 XX
 PR 04-NOV-1998; 98US-0107078.
 PR 21-NOV-1997; 97FR-0014673.
 XX
 PA (GEST) GENSET.
 XX
 PI Griffais R;
 XX
 DR WPI; 1999-357842/30.
 XX
 PS Genome sequence of Chlamydia pneumoniae
 XX
 PS Page 1429; Disclosure; 1912pp; English.
 XX
 CC Y34584-Y35879 represent the proteins encoded by all the open reading
 CC frames in the complete genome (see X91990) of Chlamydia pneumoniae.
 CC C. pneumoniae causes respiratory disease such as pneumonia and
 CC bronchitis and is thought to be a contributing factor in heart
 CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema
 CC nodosum or pharyngitis. The polypeptides encoded by the open reading

CC frames of the C. pneumoniae genome (see Y34584-Y35879) can be used in
 CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae
 CC nucleotide sequences can also be used as immunogenic compositions,
 CC especially where the vector directs the expression of a neutralising
 CC epitope of C. pneumoniae.
 XX
 SQ Sequence 94 AA;

Query Match 4.1%; Score 7; DB 20; Length 94;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 10 SLLCLVL 16
 DB 69 slclvl 75

RESULT 18
 Y31236
 ID Y31236 standard; Protein; 208 AA.
 XX
 AC Y31236;
 XX
 DT 08-NOV-1999 (first entry)
 XX
 DE Human Apo B protein N-terminal.
 XX
 KW Apo E; Apo B; hyperlipidemia; human; treatment; hepatocyte; apoprotein;
 KW Apo A1; low density lipoprotein; LDL; blood; therapy; atherosclerosis;
 KW high density lipoprotein; HDL; cholesterol; coronary heart disease;
 KW Alzheimer's disease; hypobetalipoproteinemia; dysbetalipoproteinemia.
 XX
 OS Homo sapiens.
 XX
 PN W09940789-A1.
 XX
 PD 19-AUG-1999.
 XX
 PF 28-AUG-1998; 98WO-US17908.
 XX
 PR 30-JUN-1998; 98US-0108006.
 PR 12-FEB-1998; 98US-0074497.
 XX
 PA (MINU) UNIV MINNESOTA.
 PA (YESH) UNIV YESHIVA EINSTEIN COLLEGE.
 XX
 PI Bandyopadhyay PT, Kren BT, Roy-Chowdhury J, Steer CJ;
 XX
 DR WPI; 1999-527333/44.
 DR N-PSDB; Z09525.
 XX
 PT Mutating apolipoprotein genes in hepatocytes to control cholesterol
 PT levels, e.g. for treating or preventing hyperlipidemia, particularly
 PT atherosclerosis
 XX
 PS Claim 4; Page 73; 106pp; English.
 XX
 CC This invention describes a novel method for the genetic treatment of
 CC hyperlipidemia by altering genes, in hepatocytes, for apoprotein (apo) B,
 CC E or A1. Low density lipoprotein (LDL) levels in the blood are reduced by
 CC altering an apo B gene (I) in a hepatocyte. The invention describes a
 CC method for the therapeutic and/or prophylactic method involving altering
 CC an apo E gene in hepatocytes by introducing the mutations Arg112Cys,
 CC Arg158Cys or Cys158Arg and a method for ameliorating atherosclerosis by
 CC altering the apo A1 gene in a hepatocyte so that the altered protein can
 CC dimerize. Altering expression of apo genes regulates levels of high and
 CC low density lipoprotein cholesterol. Altering expression of apo B, E and
 CC A1 genes is used to treat or prevent atherosclerosis, coronary heart
 CC disease, Alzheimer's disease, hypobetalipoproteinemia, and
 CC dysbetalipoproteinemia. This sequence represents the N-terminal fragment
 CC of the human Apo B protein described in the method of the invention.
 XX

SQ Sequence 208 AA;

Query Match 4.1%; Score 7; DB 20; Length 208;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LALPALL 22
 |||||
 Db 12 lalpell 18

RESULT 19
P70659

ID P70659 standard; Protein; 291 AA.

XX P70659;

XX 18-APR-1991 (first entry)

XX Sequence of the N-terminal portion of apolipoprotein B (apoB).

XX Drug delivery; amphipathic region; apolipoprotein.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Peptide 1..27

XX /label= hydrophobic signal

XX Protein 28..291

XX W08702061-A.

XX 09-APR-1987.

XX 02-OCT-1986; 86WO-US02074.

XX 03-OCT-1985; 85US-0783787.

XX (BIOT-) BIOTECHN RES PARTN.

XX Protter AA, Kane JP;

XX WPI; 1987-108702/15.

XX N-PSDB; N71005.

XX Drug delivery system - comprises target specific peptide having
 PT low density lipoprotein receptor-binding region bound to drug or
 PT carrier

XX Example; Fig 2; 52pp; English.

XX A compsn. for delivery of an active ingredient to a vertebrate
 CC subject comprises a target specific peptide comprising a low density
 CC lipoprotein (LDL) receptor binding region covalently bonded to a
 CC substance selected from an active ingredient and a carrier for an
 CC active ingredient is claimed. The LDL-receptor binding region may
 CC be deriv. from apolipoproteins B or E. Alternatively, alpha-helical
 CC amphipathic regions are used to bind directly to hydrophobic active
 CC ingredients. Pref the LDL-receptor region is derived from AAs 1-300
 CC of apoB-26 and contains AAs 18-24, 87-90 and 225-228, see P70658.

SQ Sequence 291 AA;

Query Match 4.1%; Score 7; DB 8; Length 291;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LALPALL 22
 |||||
 Db 12 lalpell 18

RESULT 20

Y35164

ID Y35164 standard; Protein; 352 AA.

XX Y35164;

XX 13-SEP-1999 (first entry)

XX Amino acid sequence of a Chlamydia pneumoniae protein.

XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
 KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
 KW vaccine; neutralising epitope.

XX Chlamydia pneumoniae.

XX W09927105-A2.

XX 03-JUN-1999.

XX 20-NOV-1998; 98WO-IB01890.

XX 04-NOV-1998; 98US-0107078.

XX 21-NOV-1997; 97FR-0014673.

XX (GEST) GENSET.

XX Griffais R;

XX WPI; 1999-357842/30.

XX Genome sequence of Chlamydia pneumoniae

XX Page 1030; Disclosure; 1912pp; English.

XX Y34584-Y35879 represent the proteins encoded by all the open reading
 CC frames in the complete genome (see X91990) of Chlamydia pneumoniae.
 CC C. pneumoniae causes respiratory disease such as pneumonia and
 CC bronchitis and is thought to be a contributing factor in heart
 CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema
 CC nodosum or pharyngitis. The polypeptides encoded by the open reading
 CC frames of the C. pneumoniae genome (see Y34584-Y35879) can be used in
 CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae
 CC nucleotides sequences can also be used as immunogenic compositions,
 CC especially where the vector directs the expression of a neutralising
 CC epitope of C. pneumoniae.

SQ Sequence 352 AA;

Query Match

Best Local Similarity 4.1%; Score 7; DB 20; Length 352;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 RDAETGE 45

Db 104 rdaetge 110

RESULT 21

P70661

ID P70661 standard; Protein; 435 AA.

XX P70661;

XX 18-APR-1991 (first entry)

XX Sequence of fusion protein containing apoB receptor binding and
 DE apoAI amphipathic regions.

XX Drug delivery; amphipathic region; apolipoprotein.

```
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..27
FT /label= hydrophobic signal
FT Protein 28..435
XX
XX WO8702061-A.
XX
XX 09-APR-1987.
XX
XX 02-OCT-1986; 86WO-US02074.
XX
XX 03-OCT-1985; 85US-0783787.
XX
XX (BIOT-) BIOTECHN RES PARTN.
XX
XX Protter AA, Kane JP;
XX
XX WPI: 1987-108702/15.
XX
XX N-PSDB; N71008.
XX
XX Drug delivery system - comprises target specific peptide having
XX low density lipoprotein receptor-binding region bound to drug or
XX carrier
XX
XX Example; Fig 5; 52pp; English.
XX
XX A compsn. for delivery of an active ingredient to a vertebrate
XX subject comprises a target specific peptide comprising a low density
XX lipoprotein (LDL) receptor binding region covalently bonded to a
XX substance selected from an active ingredient and a carrier for an
XX active ingredient is claimed. The LDL-receptor binding region may
XX be deriv. from apolipoproteins B or E. Alternatively, alpha-helical
XX amphipathic regions are used to bind directly to hydrophobic active
XX ingredients. Pref the LDL-receptor region is derived from AAs 1-300
XX of apoB-26 and contains AAs 18-24, 87-90 and 225-228, see P70658.
XX
XX Sequence 435 AA;

Query Match 4.1%; Score 7; DB 8; Length 435;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LAIPALL 22
Db |||||
12 laipall 18

RESULT 22
P70660
ID P70660 standard; Protein; 447 AA.
XX
XX AC P70660;
XX
XX 18-APR-1991 (first entry)
XX
XX Sequence of the N-terminal portion of apolipoprotein B (apoB).
XX
XX Drug delivery; amphipathic region; apolipoprotein.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT Peptide 1..27
FT /label= hydrophobic signal
FT Protein 28..447
XX
XX WO8702061-A.
XX
XX 09-APR-1987.
XX
```

```
PF 02-OCT-1986; 86WO-US02074.
XX
XX 03-OCT-1985; 85US-0783787.
XX
XX (BIOT-) BIOTECHN RES PARTN.
XX
XX PI Protter AA, Kane JP;
XX
XX WPI: 1987-108702/15.
XX
XX N-PSDB; N71005.
XX
XX Drug delivery system - comprises target specific peptide having
XX low density lipoprotein receptor-binding region bound to drug or
XX carrier
XX
XX Example; Fig 3; 52pp; English.
XX
XX A compsn. for delivery of an active ingredient to a vertebrate
XX subject comprises a target specific peptide comprising a low density
XX lipoprotein (LDL) receptor binding region covalently bonded to a
XX substance selected from an active ingredient and a carrier for an
XX active ingredient is claimed. The LDL-receptor binding region may
XX be deriv. from apolipoproteins B or E. Alternatively, alpha-helical
XX amphipathic regions are used to bind directly to hydrophobic active
XX ingredients. Pref the LDL-receptor region is derived from AAs 1-300
XX of apoB-26 and contains AAs 18-24, 87-90 and 225-228, see P70658.
XX
XX Sequence 447 AA;

Query Match 4.1%; Score 7; DB 8; Length 447;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LAIPALL 22
Db |||||
12 laipall 18

RESULT 23
W62832
ID W62832 standard; Protein; 590 AA.
XX
XX AC W62832;
XX
XX 27-OCT-1998 (first entry)
XX
XX Gossypium hirsutum antimicrobial protein.
XX
XX antimicrobial protein; infestation; control.
XX
XX Gossypium hirsutum.
XX
XX WO9827805-A1.
XX
XX 02-JUL-1998.
XX
XX 22-DEC-1997; 97WO-AU00874.
XX
XX 20-DEC-1996; 96AU-0004275.
XX
XX (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX
XX Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
XX
XX WPI: 1998-377279/32.
XX
XX Novel anti-microbial protein from e.g. Macadamia integrifolia -
XX useful for controlling microbial infestations of plants or mammals
XX
XX Claim 1; Page 49-51; 96pp; English.
XX
XX The sequence is that of an antimicrobial protein which can
XX
```

CC be used to control microbial infestations in plants and mammalian
 CC animals.
 XX
 SQ Sequence 590 AA;

Query Match 4.1%; Score 7; DB 19; Length 590;
 Best Local Similarity 100.0%; Pred. No. 76;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 96 GEREEA 102
 |||||

Db 159 gereeea 165

RESULT 24

ID Y59207 standard; Protein; 655 AA.

AC Y59207;

DT 21-MAR-2000 (first entry)

DE B. halodurans clone C4538 rhamnogalacturonan hydrolase.

KW Rhamnogalacturonan hydrolase; RGH; glycosyl hydrolase; bacteria;
 KW pectin; textile; detergent; animal feed; cellulose fibre.

OS Bacillus halodurans.

XX WO9957255-A1.

PN 11-NOV-1999.

PD 03-MAY-1999; 99WO-DK00244.

PF 01-MAY-1998; 98DK-0000608.

PR 05-MAY-1998; 98US-0084358.

XX (NOVO) NOVO-NORDISK AS.

PI Jorgensen PL, Schnorr K, Andersen LN, Schuelein M, Outtrup H;

XX WPI; 2000-072230/06.

DR N-PSDB; 248296.

XX New rhamnogalacturonan hydrolase enzymes used in, e.g. the textile
 PT industry -

PS Claim 45; Page 110-112; 94pp; English.

XX The invention provides isolated enzymes exhibiting rhamnogalacturonan
 CC hydrolase (RGH) activity, and belonging to a glycosyl hydrolase family
 CC other than family 28. DNA sequences 248290-299 encoding RGH enzymes.
 CC Y59201-210 respectively, obtained from various bacterial strains (see
 CC descriptor for the bacterial species and clone ID), are provided. The
 CC RGH degrades rhamnogalacturonan backbone of hairy regions of pectin. The
 CC enzymes are used in the textile, detergent, animal feed and cellulose
 CC fibre processing industries. In particular, the enzymes are used to
 CC reduce the viscosity of plant material, particularly animal feed
 CC containing soy, pea, rapeseed or other material derived from Fabales or
 CC Cruciferaeae, and to increase digestibility of that food.

XX Sequence 655 AA;

Query Match 4.1%; Score 7; DB 21; Length 655;

Best Local Similarity 100.0%; Pred. No. 83;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 RDAETGE 45

Db 396 rdaetge 402

RESULT 25

ID Y09371 standard; protein; 809 AA.

XX Y09371;

DT 15-JUL-1999 (first entry)

XX Human apolipoprotein B17 protein sequence.

XX Human; apolipoprotein B17; apoB17; low density lipoprotein; LDL;
 KW heparin; blood vessel matrix; binding inhibition; atherosclerosis;
 KW apolipoprotein B.

XX Homo sapiens.

XX WO9918986-A1.

XX 22-APR-1999.

XX 09-OCT-1998; 98WO-US21345.

XX 07-NOV-1997; 97US-0966296.

XX 10-OCT-1997; 97US-0063468.

XX (UYCO) UNIV COLUMBIA NEW YORK.

XX Goldberg IJ, Pillarisetti S;

XX WPI; 1999-287883/24.

XX Inhibition of the binding of low density lipoprotein to blood vessel
 PT matrix

XX Claim 3; Fig 1; 80pp; English.

XX The present invention describes a method for inhibiting the binding of
 CC low density lipoprotein (LDL) to blood vessel matrix in a subject. The
 CC method comprises administering to the subject a substance capable of
 CC binding to or competing with the amino-terminal region of apolipoprotein
 CC B (NTAB). Also described are: (a) a method of treating atherosclerosis;
 CC (b) identifying a compound capable of ameliorating atherosclerosis; (c)
 CC a compound identified by (b); (d) kits for inhibiting the binding of LDL
 CC to blood vessel matrix; and (e) a mutant non-human organism that
 CC overproduces NTAB. The methods are useful for identifying compounds that
 CC inhibit binding of LDL to blood vessel matrix via the NTAB or compounds
 CC that compete with NTAB for binding to blood vessel matrix. These
 CC compounds, antibodies and B17 are useful for ameliorating
 CC atherosclerosis, the major cause of coronary heart disease and strokes.
 CC Competitive antagonists are useful in cases of overproduction of NTAB.
 CC The present sequence represents a specifically claimed NTAB designated
 CC apolipoprotein B17 or apoB17.

XX Sequence 809 AA;

Query Match 4.1%; Score 7; DB 20; Length 809;

Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LALPALL 22

Db 12 lalpal 18

RESULT 26

ID Y49128 standard; Protein; 1035 AA.

XX Y49128;

```

DT 07-JAN-2000 (first entry)
DE Chimeric hmGluR2/hCaR protein sequence.
XX
XX
XX G-protein fusion receptor; CaR; calcium receptor; GluR; head injury;
KW metabotropic glutamate receptor; GABABR; chimeric receptor; stroke;
KW gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
KW spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;
KW Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
XX cognitive disorder.
XX
OS Homo sapiens.
XX
PN WO9951641-A1.
XX
PD 14-OCT-1999.
XX
XX
PF 02-APR-1999; 99WO-US07333.
XX
PR 03-APR-1998; 98US-0080671.
XX
XX (NPSP-) NPS PHARM INC.
PA Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;
PI Simin RT;
XX
XX WPI: 1999-610995/52.
DR N-PSDB; Z31059.
XX
XX New G-protein fusion receptors and chimeras containing domains from
PT different receptors, used to screen for modulators, potentially useful
PT e.g. for treating or preventing stroke or Alzheimer's disease -
PS Disclosure; Fig 10; 255pp; English.
XX
XX The invention relates to G-protein fusion receptors (I) comprising:
CC (1) in the N to C direction, extracellular (ECD), transmembrane (TMD) and
CC intracellular (ICD) domains, each chosen independently from a CaR
CC (calcium receptor), GluR (metabotropic glutamate receptor) and GABABR
CC (gamma-aminobutyric acid receptor); (2) an optional linker attached to
CC the C-terminus of ICD; and (3) a G-protein (GP) linked to ICD or the
CC linker. (I), and recombinant chimeric receptors (CR) without the GP
CC component, are used to assess function of the various domains and to
CC identify compounds (e.g. allosteric modulators or antagonists) that act
CC on these domains. The modulators are potentially useful for treating or
CC preventing diseases associated with the receptors, e.g. stroke, head or
CC spinal cord injury, epilepsy, ischaemia, hypoglycaemia, anoxia,
CC Alzheimer's disease, hyperparathyroidism, osteoporosis, cognitive
CC disorders and depression. Nucleic acid (II) that encodes (I) is used:
CC (1) for recombinant production of corresponding proteins; and (2) to
CC produce cells used in screening for modulators. Use of CaR and mGluR
CC domains allows presentation of GABABR domains, to a binding agent, in a
CC form more like the natural domain structure compared with use of
CC incomplete receptors, lacking one or more domains. By shuffling different
CC domains, agents can be identified that affect particular domains of a
CC receptor.
XX
XX Sequence 1035 AA;
SQ
Query Match 4.1%; Score 7; DB 20; Length 1035;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 16 LALPALL 22
Db 5 lalball 11
RESULT 27
Y49129
ID Y49129 standard; Protein; 1394 AA.
XX
AC Y49129;
XX
XX 07-JAN-2000 (first entry)
XX
XX pmGluR2/CaR*Galphaq15 fusion construct protein sequence.
XX
XX G-protein fusion receptor; CaR; calcium receptor; GluR; head injury;
KW metabotropic glutamate receptor; GABABR; chimeric receptor; stroke;
KW gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
KW spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;
KW Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
XX cognitive disorder.
XX
XX Homo sapiens.
XX
XX WO9951641-A1.
XX
XX 14-OCT-1999.
XX
XX
XX 02-APR-1999; 99WO-US07333.
XX
XX 03-APR-1998; 98US-0080671.
XX
XX (NPSP-) NPS PHARM INC.
PA Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;
PI Simin RT;
XX
XX WPI: 1999-610995/52.
DR N-PSDB; Z31060.
XX
XX New G-protein fusion receptors and chimeras containing domains from
PT different receptors, used to screen for modulators, potentially useful
PT e.g. for treating or preventing stroke or Alzheimer's disease -
PS Example 1; Fig 12; 255pp; English.
XX
XX The invention relates to G-protein fusion receptors (I) comprising:
CC (1) in the N to C direction, extracellular (ECD), transmembrane (TMD) and
CC intracellular (ICD) domains, each chosen independently from a CaR
CC (calcium receptor), GluR (metabotropic glutamate receptor) and GABABR
CC (gamma-aminobutyric acid receptor); (2) an optional linker attached to
CC the C-terminus of ICD; and (3) a G-protein (GP) linked to ICD or the
CC linker. (I), and recombinant chimeric receptors (CR) without the GP
CC component, are used to assess function of the various domains and to
CC identify compounds (e.g. allosteric modulators or antagonists) that act
CC on these domains. The modulators are potentially useful for treating or
CC preventing diseases associated with the receptors, e.g. stroke, head or
CC spinal cord injury, epilepsy, ischaemia, hypoglycaemia, anoxia,
CC Alzheimer's disease, hyperparathyroidism, osteoporosis, cognitive
CC disorders and depression. Nucleic acid (II) that encodes (I) is used:
CC (1) for recombinant production of corresponding proteins; and (2) to
CC produce cells used in screening for modulators. Use of CaR and mGluR
CC domains allows presentation of GABABR domains, to a binding agent, in a
CC form more like the natural domain structure compared with use of
CC incomplete receptors, lacking one or more domains. By shuffling different
CC domains, agents can be identified that affect particular domains of a
CC receptor.
XX
XX Sequence 1394 AA;
SQ
Query Match 4.1%; Score 7; DB 20; Length 1394;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 16 LALPALL 22
Db 5 lalball 11
RESULT 28
Y49134

```

ID Y49134 standard; Protein; 1397 AA.
 XX Y49134;
 AC
 XX
 DT 07-JAN-2000 (first entry)
 XX
 XX pmGluR2/Car*Galphai5+3Aa linker fusion construct protein sequence.
 DE
 XX
 XX G-protein fusion receptor; CaR; calcium receptor; GluR; head injury;
 KW metabotropic glutamate receptor; GABABR; chimeric receptor; stroke;
 KW gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
 KW spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;
 KW Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
 KW cognitive disorder.
 KW
 XX Homo sapiens.
 OS
 XX
 XX WO951641-A1.
 PN
 XX
 XX 14-OCT-1999.
 PD
 XX
 XX 02-APR-1999; 99WO-US07333.
 PF
 XX
 XX 03-APR-1998; 98US-0080671.
 PR
 XX
 XX (NPSP-) NPS PHARM INC.
 PA
 XX
 XX Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;
 PI Simin RT;
 PI
 XX
 XX WPI; 1999-610995/52.
 DR
 XX
 XX N-PSDB; 231065.
 DR
 XX
 XX New G-protein fusion receptors and chimeras containing domains from
 PT different receptors, used to screen for modulators, potentially useful
 PT e.g. for treating or preventing stroke or Alzheimer's disease -
 PT
 XX
 XX Example 1; Fig 12; 255pp; English.
 PS
 XX
 CC The invention relates to G-protein fusion receptors (I) comprising:
 CC (1) in the N to C direction, extracellular (ECD), transmembrane (TMD) and
 CC intracellular (ICD) domains, each chosen independently from a CaR
 CC (calcium receptor), GluR (metabotropic glutamate receptor) and GABABR
 CC (gamma-aminobutyric acid receptor); (2) an optional linker attached to
 CC the C-terminus of ICD; and (3) a G-protein (GP) linked to ICD or the
 CC linker. (I), and recombinant chimeric receptors (CR) without the GP
 CC component, are used to assess function of the various domains and to
 CC identify compounds (e.g. allosteric modulators or antagonists) that act
 CC on these domains. The modulators are potentially useful for treating or
 CC preventing diseases associated with the receptors, e.g. stroke, head or
 CC spinal cord injury, epilepsy, ischaemia, hypoglycaemia, anoxia,
 CC Alzheimer's disease, hyperparathyroidism, osteoporosis, cognitive
 CC disorders and depression. Nucleic acid (II) that encodes (I) is used;
 CC (1) for recombinant production of corresponding proteins; and (2) to
 CC produce cells used in screening for modulators. Use of CaR and mGluR
 CC domains allows presentation of GABABR domains, to a binding agent, in a
 CC form more like the natural domain structure compared with use of
 CC incomplete receptors, lacking one or more domains. By shuffling different
 CC domains, agents can be identified that affect particular domains of a
 CC receptor.
 XX
 SQ Sequence 1397 AA;

Query Match 4.1%; Score 7; DB 20; Length 1397;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LALPALL 22
 |||||
 Db 5 lalpaill 11

RESULT 29
 W22604
 ID W22604 standard; Protein; 1611 AA.
 XX
 AC W22604;
 XX
 DT 27-FEB-1998 (first entry)
 XX
 XX Ty lactone synthase ORF4 protein.
 DE
 XX Ty lactone synthase gene cluster; tyIG gene; multifunctional protein;
 KW polyketide; ty lactone synthesis; antibiotic; tylosin.
 KW
 XX Streptomyces fradiae.
 OS
 XX
 XX Key Location/Qualifiers
 FT Domain 36..459
 FT /note= "ketosynthase domain, KS6"
 FT Domain 582..910
 FT /note= "acyltransferase domain, AT6"
 FT Domain 1155..1335
 FT /note= "ketoreductase domain, KR6"
 FT Domain 1421..1504
 FT /note= "acyl carrier protein domain, ACP6"
 FT
 XX EP791655-A2.
 PN
 XX
 PD 27-AUG-1997.
 PD
 XX
 PF 19-FEB-1997; 97EP-0301056.
 PF
 XX
 PR 22-FEB-1996; 96US-0012078.
 PR
 XX
 XX (ELIL) LILLY & CO ELI.
 PA
 XX
 PI Dehoff BS, Kuhstoss SA, Rostock PR, Sutton KL;
 PI
 XX WPI; 1997-418046/39.
 DR
 XX N-PSDB; T80413.
 DR
 XX
 CC DNA encoding Streptomyces fradiae ty lactone synthase domain - for
 CC production of tylosin-related polyketide compounds
 CC
 PS Claim 9; Pages 98-103; 220pp; English.
 XX
 CC W22601-W22605 represent proteins encoded by the ty lactone synthase gene
 CC cluster of the invention. The gene cluster is also referred to as the
 CC tyIG gene, and was isolated from Streptomyces fradiae. These sequences
 CC are multifunctional proteins which direct the synthesis of the polyketide
 CC ty lactone, isolated from Streptomyces fradiae. Ty lactone is the basic
 CC building block of the antibiotic tylosin. The DNA sequence can be
 CC modified so as to alter the type of carboxylic acids incorporated, the
 CC number of carboxylic acids incorporated and/or the post-condensation
 CC reactions performed, thereby resulting in novel tylosin-related
 CC polyketides.
 CC
 XX
 SQ Sequence 1611 AA;

Query Match 4.1%; Score 7; DB 18; Length 1611;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 98 REEEARA 104
 |||||
 Db 1516 reeeara 1522

RESULT 30
 R06341
 ID R06341 standard; protein; 1671 AA.
 XX
 AC R06341;

```

XX 14-DEC-1990 (first entry)
DT Signal peptide and first 1643 AAs of mature apoB protein.
XX
DE
XX
XX Atherosclerosis; lipid binding peptide; LBP; phospholipid;
KW lecithin cholesterol acyl transferase; L-CAT; apo A1;
KW Chinese hamster ovaries;
XX
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT Peptide 1..27
FT /label=Signal peptide.
FT Protein 28..1671
FT /label=Mature peptide.
XX
XX US4943527-A.
XX
XX 24-JUL-1990.
XX
XX 27-FEB-1986; 86US-0834300.
XX
XX 27-FEB-1986; 86US-0834300.
PR 04-OCT-1985; 85US-0784418.
PR 04-DEC-1985; 85US-0804692.
XX
XX (CALB-) CALIFORNIA BIOTECHN.
XX
XX Protter AA, Vigne JL, Mallory JB, Talmadge KD, Kane JP;
PI WPI; 1990-246622/32.
XX
XX N-PSDB; Q05590.
XX
XX Prodn. of purified lipid-binding peptide - capable of binding to
PT phospholipid(s), useful in e.g. atherosclerosis.
XX
XX Disclosure; ; pp; English.
XX
XX Lipid-binding peptides in conjunction with phospholipids are
CC effective in reversing cholesterol transport, and may enhance
CC lecithin cholesterol acetyl transferase, lowering the rate of
CC plaque formation and atherosclerosis.
XX
XX Sequence 1671 AA;

Query Match 4.1%; Score 7; DB 11; Length 1671;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LALPALL 22
Db 12 laipall 18

RESULT 31
P70647
ID P70647 standard; protein; 2721 AA.
AC
AC P70647;
XX
XX 17-APR-1991 (first entry)
DT
XX
XX Sequence of N-terminal apolipoprotein B (apoB).
DE
XX
XX Lipid-binding peptide; apolipoprotein purification.
KW
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH Peptide 87..167
FT /label= Signal

```

```

FT Protein 168..8247
XX
XX W08702062-A.
XX
XX 09-APR-1987.
PD
XX
XX 02-OCT-1986; 86WO-US02075.
PF
XX
XX 04-DEC-1985; 85US-0804692.
PR
XX 04-OCT-1985; 85US-0784418.
PR
XX 27-FEB-1986; 86US-0834300.
XX
XX (BIOT-) BIOTECHN RES PARTN.
PA
XX
XX Protter AA, Vigne JL, Mallory JB, Talmadge KD;
PI WPI; 1987-108703/15.
XX
XX N-PSDB; N70996.
XX
XX Purified lipid-binding peptide prodn. - by expression of
PT recombinant system and adding lipid to obtd. peptide mixt. to
PT form lipo-peptide complex
XX
XX Claim 50; pp73-85; 123pp; English.
PS
XX
XX ApoB is a lipid-binding peptide adapted to bind to phospholipids at
CC one or more amphipatic alpha-helical peptide regions. ApoB gene is
CC pref. expressed in Chinese hamster ovary (CHO) cells and the gene is
CC introduced in an expression vector having a regulatable promoter
CC derived from the human metallothionein II gene.
XX
XX Sequence 2721 AA;

Query Match 4.1%; Score 7; DB 8; Length 2721;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LALPALL 22
Db 12 laipall 18

RESULT 32
Y60941
ID Y60941 standard; Peptide; 7 AA.
XX
XX AC Y60941;
AC
XX
XX 02-MAR-2000 (first entry)
DT
XX
XX Cadherin-5 cell adhesion recognition cyclic peptide SEQ ID NO:961.
DE
XX
XX Modulation; nonclassical cadherin mediated cell adhesion; CAR;
KW inhibition; cadherin extracellular domain; cell adhesion recognition;
KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;
KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
KW cadherin related neuronal receptor; LI-cadherin; protocadherin;
KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
KW neurotoxic disease; cyclic.
XX
XX Synthetic.
OS
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Modified-site 1..7
FT /note= "the terminal residues are condensed with each
FT other to form a cyclic peptide"
XX
XX W09957149-A2.
XX
XX 11-NOV-1999.
PD

```



```

XX PF 05-MAY-1999; 99WO-CA00363.
XX PR 05-MAY-1998; 98US-0073040.
XX PR 06-NOV-1998; 98US-0187859.
XX PR 20-JAN-1999; 99US-0234395.
XX PR 08-MAR-1999; 99US-0264516.
XX PA (ADHE-) ADHEREX TECHNOLOGIES INC.
XX PI Blaschuk OW, Gour BJ, Byers S;
XX DR WPI; 2000-038791/03.
XX PT New cadherin modulating agents, used for modulating nonclassical
XX PT cadherin-mediated functions for treating e.g. cancers, obesity,
XX PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological
XX PS disease -
XX PS Claim 24; Page 163; 252pp; English.
XX CC The present invention describes cadherin modulating agents (MA)
XX CC comprising peptides which comprise a nonclassical cadherin cell adhesion
XX CC recognition (CAR) sequence. The MAs can be used for modulating
XX CC nonclassical cadherin-mediated functions. They can be used for e.g.
XX CC inhibiting adhesion of nonclassical-cadherin expressing cells in a
XX CC mammal, enhancing delivery of a drug through the skin of a mammal,
XX CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in
XX CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
XX CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
XX CC expressing cell, preventing or treating obesity in a mammal, stimulating
XX CC blood vessel regression in a mammal, enhancing drug delivery to the
XX CC central nervous system, treating a demyelinating neurological disease,
XX CC increasing vasopermeability in a mammal, enhancing adhesion of
XX CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in
XX CC a mammal, or preventing pregnancy in a mammal. They can also be used for
XX CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing
XX CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a
XX CC mammal. They can also be used for treating e.g. psoriasis, arthritis,
XX CC age-related macular degeneration, multiple sclerosis and diabetes. The
XX CC products can also be used for detection and diagnosis and in
XX CC bioreactors. Y60592 to Y64572 represent specifically claimed peptides,
XX CC and Y64573 to Y64643 and Z33183 to Z33186 represent sequences used in the
XX CC exemplification of the present invention.
XX SQ Sequence 7 AA;

Query Match 3.5%; Score 6; DB 21; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.le+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 DAETGE 45
Db | | | | |
2 daetge 7

RESULT 33
Y60942
ID Y60942 standard; Peptide: 8 AA.
XX AC Y60942;
XX 02-MAR-2000 (first entry)
XX Cadherin-5 cell adhesion recognition cyclic peptide SEQ ID NO:962.
XX Modulation; nonclassical cadherin mediated cell adhesion; CAR;
XX inhibition; cadherin extracellular domain; cell adhesion recognition;
XX OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;
XX cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
XX cadherin related neuronal receptor; LI-cadherin; protocadherin;
XX desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;

```

```

KW KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
XX XX neurological disease; cyclic.
OS Synthetic.
OS Homo sapiens.
XX Key Location/Qualifiers
FH Modified-site 1.8
FT /note= "the terminal residues are condensed with each
FT other to form a cyclic peptide"
XX WO9957149-A2.
XX 11-NOV-1999.
XX 05-MAY-1999; 99WO-CA00363.
XX 05-MAY-1998; 98US-0073040.
XX 06-NOV-1998; 98US-0187859.
XX 20-JAN-1999; 99US-0234395.
XX 08-MAR-1999; 99US-0264516.
XX (ADHE-) ADHEREX TECHNOLOGIES INC.
XX Blaschuk OW, Gour BJ, Byers S;
XX WPI; 2000-038791/03.
XX New cadherin modulating agents, used for modulating nonclassical
XX PT cadherin-mediated functions for treating e.g. cancers, obesity,
XX PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological
XX PT disease -
XX Claim 24; Page 163; 252pp; English.
XX The present invention describes cadherin modulating agents (MA)
XX CC comprising peptides which comprise a nonclassical cadherin cell adhesion
XX CC recognition (CAR) sequence. The MAs can be used for modulating
XX CC nonclassical cadherin-mediated functions. They can be used for e.g.
XX CC inhibiting adhesion of nonclassical-cadherin expressing cells in a
XX CC mammal, enhancing delivery of a drug through the skin of a mammal,
XX CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in
XX CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
XX CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
XX CC expressing cell, preventing or treating obesity in a mammal, stimulating
XX CC blood vessel regression in a mammal, enhancing drug delivery to the
XX CC central nervous system, treating a demyelinating neurological disease,
XX CC increasing vasopermeability in a mammal, enhancing adhesion of
XX CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in
XX CC a mammal, or preventing pregnancy in a mammal. They can also be used for
XX CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing
XX CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a
XX CC mammal. They can also be used for treating e.g. psoriasis, arthritis,
XX CC age-related macular degeneration, multiple sclerosis and diabetes. The
XX CC products can also be used for detection and diagnosis and in
XX CC bioreactors. Y60592 to Y64572 represent specifically claimed peptides,
XX CC and Y64573 to Y64643 and Z33183 to Z33186 represent sequences used in the
XX CC exemplification of the present invention.
XX SQ Sequence 8 AA;

Query Match 3.5%; Score 6; DB 21; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.le+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 DAETGE 45
Db | | | | |
3 daetge 8

RESULT 34
W10275

```

ID W10275 standard; peptide; 9 AA.
 XX
 AC W10275;
 XX
 DT 21-AUG-1997 (first entry)
 XX
 DE Antiphospholipid analogue 3B6 which binds to B cells.
 XX
 KW aPL; antibody mediated pathology; systemic lupus erythematosus;
 KW antiphospholipid antibody syndrome; stroke; foetal loss; thrombosis.
 XX
 OS Synthetic.
 XX
 PN W09640197-A1.
 XX
 PD 19-DEC-1996.
 XX
 PF 06-JUN-1996; 96WO-US09976.
 XX
 PR 07-JUN-1995; 95US-0482651.
 XX
 PA (LJOL-) LA JOLLA PHARM CO.
 XX
 PI Jones DS, Marquis DM, Victoria EJ, Yu L;
 XX
 DR WPI; 1997-051886/05.
 XX
 XX New anti-phospholipid analogues - used to develop prods. for the
 PT diagnosis and treatment of aPL antibody-mediated pathologies such as
 PT systemic lupus erythematosus
 XX
 PS Claim 4; Page 84; 118pp; English.
 XX
 CC The present sequence is an antiphospholipid (aPL) analogue that binds
 CC specifically to B cells to which an aPL epitope binds. Peptides
 CC are derived from screening random peptide phage libraries to identify
 CC the sequences which best mimic the epitopes recognised by aPL
 CC antibodies. A method has been produced for identifying analogues of
 CC key epitopes recognised by aPL antibodies in patients suffering from
 CC PAPS, APS and other aPL antibody-mediated diseases. The products and
 CC methods can be used for the diagnosis and treatment of aPL antibody-
 CC mediated pathologies such as systemic lupus erythematosus,
 CC antiphospholipid antibody syndrome, recurrent stroke, recurrent foetal
 CC loss or thrombosis.
 XX
 SQ Sequence 9 AA;
 Query Match 3.5%; Score 6; DB 18; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.1e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 13 CLVLAL 18
 Db 1 clvlal 6
 RESULT 35
 W67101
 ID W67101 standard; peptide; 9 AA.
 XX
 AC W67101;
 XX
 DT 18-DEC-1998 (first entry)
 XX
 DE Antibody ACA-6501 epitope isolated from a phage library screen.
 XX
 KW antiphospholipid; aPL; analogue; stroke; foetal loss; B cell; epitope;
 KW thrombosis; antibody ACA-6501; conjugate; phage library screen;
 KW biopanning; micropanning.
 XX
 OS Synthetic.
 XX

PN W09746251-A1.
 XX
 PD 11-DEC-1997.
 XX
 PF 06-JUN-1997; 97WO-US10075.
 XX
 PR 05-DEC-1996; 96US-0760508.
 PR 06-JUN-1996; 96US-0660092.
 XX
 PA (LJOL-) LA JOLLA PHARM CO.
 XX
 PI Jones DS, Marquis DM, Victoria EJ, Yu L;
 XX
 DR WPI; 1998-051902/05.
 XX
 PT New anti-phospholipid analogues and conjugates containing them - are
 PT useful in treatment of disorders mediated by antiphospholipid
 PT antibody, such as stroke or foetal loss
 XX
 PS Disclosure; Page 26; 155pp; English.
 XX
 CC The invention relates to antiphospholipid (aPL) analogues that bind
 CC specifically to B cells to which an aPL epitope binds. Optimised
 CC analogues lack T cell epitope(s) and are useful as conjugates for
 CC treating aPL antibody-mediated diseases. Conjugates comprising aPL
 CC analogues and non-immunogenic valency platform molecules are provided as
 CC are novel non-immunogenic valency platform molecules and linkers. Methods
 CC of preparing and identifying the analogues, methods of treatment and
 CC methods and compositions for preparing conjugates of the analogues and
 CC diagnostic immunoassays for aPL antibodies are disclosed. The conjugates
 CC are useful for treatment of aPL antibody-mediated diseases, such as
 CC stroke, foetal loss, antiphospholipid antibody syndrome or thrombosis.
 CC Antibody ACA-6501, from a patient with a history of recurrent stroke,
 CC foetal loss, lupus and aortic valve replacements, was used in phage
 CC library screens. Sequences W67097-W67132 represent epitopes that
 CC were obtained in phage that micropanned.
 XX
 SQ Sequence 9 AA;
 Query Match 3.5%; Score 6; DB 19; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.1e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 13 CLVLAL 18
 Db 1 clvlal 6
 RESULT 36
 W46066
 ID W46066 standard; peptide; 9 AA.
 XX
 AC W46066;
 XX
 DT 18-DEC-1998 (first entry)
 XX
 DE Antiphospholipid analogue.
 XX
 KW antiphospholipid; aPL; analogue; stroke; foetal loss; B cell; epitope;
 KW thrombosis; antibody; conjugate.
 XX
 OS Synthetic.
 XX
 PN W09746251-A1.
 XX
 PD 11-DEC-1997.
 XX
 PF 06-JUN-1997; 97WO-US10075.
 XX
 PR 05-DEC-1996; 96US-0760508.
 PR 06-JUN-1996; 96US-0660092.
 XX

PA (LJOL-) LA JOLLA PHARM CO.
 XX Jones DS, Marquis DM, Victoria EJ, Yu L;
 XX WPI; 1998-051902/05.
 DR
 XX New anti-phospholipid analogues and conjugates containing them - are
 PT useful in treatment of disorders mediated by antiphospholipid
 PT antibody, such as stroke or foetal loss
 XX
 XX Claim 4; 112; 155pp; English.
 PS
 XX The invention relates to antiphospholipid (aPL) analogues that bind
 CC specifically to B cells to which an aPL epitope binds. Optimised
 CC analogues lack T cell epitope(s) and are useful as conjugates for
 CC treating aPL antibody-mediated diseases. Conjugates comprising aPL
 CC analogues and non-immunogenic valency platform molecules are provided as
 CC are novel non-immunogenic valency platform molecules and linkers. Methods
 CC of preparing and identifying the analogues, methods of treatment and
 CC methods and compositions for preparing conjugates of the analogues and
 CC diagnostic immunoassays for aPL antibodies are disclosed. The conjugates
 CC are useful for treatment of aPL antibody-mediated diseases, such as
 CC stroke, foetal loss, antiphospholipid antibody syndrome or thrombosis.
 CC The present sequence represents a specifically claimed antiphospholipid
 CC analogue.
 XX
 XX Sequence 9 AA;
 SQ

Query Match 3.5%; Score 6; DB 19; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.1e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 13 CIVIAL 18
 Db 1 civil 6
 |||||

RESULT 37
 Y60943
 ID Y60943 standard; Peptide; 9 AA.
 AC
 XX Y60943;
 XX
 XX 02-MAR-2000 (first entry)
 DT
 XX Cadherin-5 cell adhesion recognition cyclic peptide SEQ ID NO:963.
 DE
 XX Modulation; nonclassical cadherin mediated cell adhesion; CAR;
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;
 KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
 KW neurological disease; cyclic.
 XX
 XX Synthetic.
 OS
 XX Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Modified-site 1..9
 FT /note= "the terminal residues are condensed with each
 FT other to form a cyclic peptide"
 XX
 XX WO9957149-A2.
 FN
 XX 11-NOV-1999.
 PD
 XX 05-MAY-1999; 99WO-CA00363.
 PF
 XX 05-MAY-1998; 98US-0073040.
 PR
 XX 06-NOV-1998; 98US-0187859.
 PR

PR 20-JAN-1999; 99US-0234395.
 PR 08-MAR-1999; 99US-0264516.
 XX (ADHE-) ADHEREX TECHNOLOGIES INC.
 PA Blaschuk OW, Gour BJ, Byers S;
 PI WPI; 2000-038791/03.
 XX
 XX New cadherin modulating agents, used for modulating nonclassical
 PT cadherin-mediated functions for treating e.g. cancers, obesity,
 PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological
 PT disease -
 XX
 XX Claim 24; Page 163; 252pp; English.
 PS
 XX The present invention describes cadherin modulating agents (MA)
 CC comprising peptides which comprise a nonclassical cadherin cell
 CC recognition (CAR) sequence. The MAs can be used for modulating
 CC nonclassical cadherin-mediated functions. They can be used for e.g.
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a
 CC mammal, enhancing delivery of a drug through the skin of a mammal,
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
 CC expressing cell, preventing or treating obesity in a mammal, stimulating
 CC blood vessel regression in a mammal, enhancing drug delivery to the
 CC central nervous system, treating a demyelinating neurological disease,
 CC increasing vasopermeability in a mammal, enhancing adhesion of
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing
 CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a
 CC mammal. They can also be used for treating e.g. psoriasis, arthritis,
 CC age-related macular degeneration, multiple sclerosis and diabetes. The
 CC products can also be used for detection and diagnosis and in
 CC bioreactors. Y60592 to Y64572 represent specifically claimed peptides,
 CC and Y64573 to Y64643 and Z33183 to Z33186 represent sequences used in the
 CC exemplification of the present invention.
 XX
 XX Sequence 9 AA;
 SQ

Query Match 3.5%; Score 6; DB 21; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2.1e+05;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 40 DAETGE 45
 Db 4 daetge 9
 |||||

RESULT 38
 Y60946
 ID Y60946 standard; Peptide; 10 AA.
 AC Y60946;
 XX
 XX 02-MAR-2000 (first entry)
 DT
 XX Cadherin-5 cell adhesion recognition cyclic peptide SEQ ID NO:966.
 DE
 XX Modulation; nonclassical cadherin mediated cell adhesion; CAR;
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;
 KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
 KW neurological disease; cyclic.
 XX
 XX Synthetic.
 OS
 XX Homo sapiens.

```

XX FH Key Location/Qualifiers
FT Modified-site 1..10
FT /note= "the terminal residues are condensed with each
XX other to form a cyclic peptide"
XX
XX WO9957149-A2.
XX
XX 11-NOV-1999.
XX
XX 05-MAY-1999; 99WO-CA00363.
XX
XX 05-MAY-1998; 98US-0073040.
XX 06-NOV-1998; 98US-0187859.
XX 20-JAN-1999; 99US-0234395.
XX 08-MAR-1999; 99US-0264516.
XX
XX (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
XX Blaschuk OW, Gour BJ, Byers S;
XX
XX WPI: 2000-038791/03.
XX
XX New cadherin modulating agents, used for modulating nonclassical
XX cadherin-mediated functions for treating e.g. cancers, obesity,
XX rheumatoid arthritis, multiple sclerosis, diabetes or a neurological
XX disease
XX
XX Claim 24; Page 163; 252pp; English.
XX
XX The present invention describes cadherin modulating agents (MA)
XX comprising peptides which comprise a nonclassical cadherin cell
XX recognition (CAR) sequence. The MAs can be used for modulating
XX nonclassical cadherin-mediated functions. They can be used for e.g.
XX inhibiting adhesion of nonclassical-cadherin expressing cells in a
XX mammal, enhancing delivery of a drug through the skin of a mammal,
XX enhancing delivery of a drug to a tumour in a mammal, treating cancer in
XX a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
XX angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
XX expressing cell, preventing or treating obesity in a mammal, stimulating
XX blood vessel regression in a mammal, enhancing drug delivery to the
XX central nervous system, treating a demyelinating neurological disease,
XX increasing vasopermeability in a mammal, enhancing adhesion of
XX nonclassical cadherin-expressing cells, inhibiting synaptic stability in
XX a mammal, or preventing pregnancy in a mammal. They can also be used for
XX e.g. enhancing or directing neurite outgrowth, facilitating wound healing
XX or reducing scar tissue, or enhancing adhesion of foreign tissue in a
XX mammal. They can also be used for treating e.g. psoriasis, arthritis,
XX age-related macular degeneration, multiple sclerosis and diabetes. The
XX products can also be used for detection and diagnosis and in
XX bioreactors. Y60592 to Y64572 represent specifically claimed peptides,
XX and Y64573 to Y64643 and Z33183 to Z33186 represent sequences used in the
XX exemplification of the present invention.
XX
XX Sequence 10 AA;
XX
XX Query Match 3.5%; Score 6; DB 21; Length 10;
XX Best Local Similarity 100.0%; Pred. No. 21;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 40 DAETGE 45
XX |
XX 5 daetge 10
XX
XX RESULT 39
XX Y60949
XX ID Y60949 standard; Peptide; 11 AA.
XX
XX AC Y60949;
XX
XX DT 02-MAR-2000 (first entry)

```

```

XX DE XX Cadherin-5 cell adhesion recognition cyclic peptide SEQ ID NO:969.
XX
XX KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;
XX inhibition; cadherin extracellular domain; cell adhesion recognition;
XX OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;
XX KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
XX KW cadherin related neuronal receptor; LI-cadherin; protocadherin;
XX KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
XX KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
XX KW neurological disease; cyclic.
XX
XX OS Synthetic.
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
FT Modified-site 1..11
FT /note= "the terminal residues are condensed with each
FT other to form a cyclic peptide"
FT
FT WO9957149-A2.
FT
FT 11-NOV-1999.
FT
FT 05-MAY-1999; 99WO-CA00363.
FT
FT 05-MAY-1998; 98US-0073040.
FT 06-NOV-1998; 98US-0187859.
FT 20-JAN-1999; 99US-0234395.
FT 08-MAR-1999; 99US-0264516.
FT
FT (ADHE-) ADHEREX TECHNOLOGIES INC.
FT
FT Blaschuk OW, Gour BJ, Byers S;
FT
FT WPI: 2000-038791/03.
FT
FT New cadherin modulating agents, used for modulating nonclassical
FT cadherin-mediated functions for treating e.g. cancers, obesity,
FT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological
FT disease
FT
FT Claim 24; Page 164; 252pp; English.
FT
FT The present invention describes cadherin modulating agents (MA)
FT comprising peptides which comprise a nonclassical cadherin cell
FT recognition (CAR) sequence. The MAs can be used for modulating
FT nonclassical cadherin-mediated functions. They can be used for e.g.
FT inhibiting adhesion of nonclassical-cadherin expressing cells in a
FT mammal, enhancing delivery of a drug through the skin of a mammal,
FT enhancing delivery of a drug to a tumour in a mammal, treating cancer in
FT a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
FT angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
FT expressing cell, preventing or treating obesity in a mammal, stimulating
FT blood vessel regression in a mammal, enhancing drug delivery to the
FT central nervous system, treating a demyelinating neurological disease,
FT increasing vasopermeability in a mammal, enhancing adhesion of
FT nonclassical cadherin-expressing cells, inhibiting synaptic stability in
FT a mammal, or preventing pregnancy in a mammal. They can also be used for
FT e.g. enhancing or directing neurite outgrowth, facilitating wound healing
FT or reducing scar tissue, or enhancing adhesion of foreign tissue in a
FT mammal. They can also be used for treating e.g. psoriasis, arthritis,
FT age-related macular degeneration, multiple sclerosis and diabetes. The
FT products can also be used for detection and diagnosis and in
FT bioreactors. Y60592 to Y64572 represent specifically claimed peptides,
FT and Y64573 to Y64643 and Z33183 to Z33186 represent sequences used in the
FT exemplification of the present invention.
FT
FT Sequence 11 AA;
FT
FT Query Match 3.5%; Score 6; DB 21; Length 11;
FT Best Local Similarity 100.0%; Pred. No. 23;
FT
FT QY 40 DAETGE 45
FT |
FT 5 daetge 10
FT
FT RESULT 39
FT Y60949
FT ID Y60949 standard; Peptide; 11 AA.
FT
FT AC Y60949;
FT
FT DT 02-MAR-2000 (first entry)

```

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 40 DAETGE 45

Db 6 daetge 11

RESULT 40

ID W30936

XX W30936 standard; peptide; 13 AA.

AC W30936;

XX 08-JAN-1998 (first entry)

XX Birch pollen allergen T cell epitope.

XX Cofactor-independent phosphoglycerate mutase; PGM-i; E.C. 5.4.21;

KW birch; pollen; allergy; plant allergen; panallergen; B cell;

KW T cell; epitope; immunotherapy; detection; diagnosis;

KW hay fever; conserved.

XX Betula verrucosa.

OS W09705258-A2.

XX 13-FEB-1997.

XX 02-AUG-1996; 96WO-AT00141.

XX 02-AUG-1995; 95AT-0001320.

XX (BIOM-) BIOMAY PRODN & HANDELS GMBH.

PI Breitenbach M, Ebner C, Engel E, Ferreira F, Jilek A;

PI Kraft D, Richter K, Rheinberger H;

XX WPI; 1997-145695/13.

XX New recombinant DNA encoding plant phosphoglycerate mutase or its

PT antigenic epitope(s) - useful for diagnosis or treatment of

PT allergies to pollen and plant-derived foods

XX Disclosure; Fig 4; 160pp; German.

XX W30936-W30949 are T cell epitopes of a birch pollen co-factor-independent

CC phosphoglycerate mutase (PGM-i) allergen. PGM-i is a highly conserved

CC plant allergen (panallergen) which can cause cross-reactivity in

CC patients allergic to pollen and plant-derived foods. PGM-i and it's B

CC cell and T cell epitopes can be used for the in vitro detection of

CC allergy against PGM-i, by measuring serum IgE or a cellular reaction.

CC They can also be used in immunotherapy and will not cause an autoimmune

CC response because PGM-i is significantly different from the human

CC enzyme, which is co-factor dependent.

XX Sequence 13 AA;

Query Match 3.5%; Score 6; DB 18; Length 13;

Best Local Similarity 100.0%; Pred. No. 27;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 VAETPT 35

Db 6 vaetpt 11

RESULT 41

ID W31042

XX W31042 standard; peptide; 16 AA.

XX AC W31042;

XX

09-JAN-1998 (first entry)

XX Mugwort pollen allergen T cell epitope.

XX Cofactor-independent phosphoglycerate mutase; PGM-i; E.C. 5.4.21;

KW Timothy grass; pollen; allergy; plant allergen; panallergen; B cell;

KW T cell; epitope; immunotherapy; detection; diagnosis; hay fever;

KW conserved.

XX Artemisia vulgaris.

OS W09705258-A2.

XX 13-FEB-1997.

XX 02-AUG-1996; 96WO-AT00141.

XX 02-AUG-1995; 95AT-0001320.

XX (BIOM-) BIOMAY PRODN & HANDELS GMBH.

PI Breitenbach M, Ebner C, Engel E, Ferreira F, Jilek A;

PI Kraft D, Richter K, Rheinberger H;

XX WPI; 1997-145695/13.

XX New recombinant DNA encoding plant phosphoglycerate mutase or its

PT antigenic epitope(s) - useful for diagnosis or treatment of

PT allergies to pollen and plant-derived foods

XX Disclosure; Fig 12a; 160pp; German.

XX W31041-W31050 are T cell epitopes of mugwort pollen co-factor-

CC independent phosphoglycerate mutase (PGM-i) isoform Art6. PGM-i is

CC a highly conserved plant allergen (panallergen) which can cause

CC cross-reactivity in patients allergic to pollen and plant-derived

CC foods. PGM-i and it's B cell and T cell epitopes can be used for the

CC in vitro detection of allergy against PGM-i, by measuring serum IgE

CC or a cellular reaction. They can also be used in immunotherapy and

CC will not cause an autoimmune response because PGM-i is significantly

CC different from the human enzyme, which is co-factor dependent.

XX Sequence 16 AA;

Query Match 3.5%; Score 6; DB 18; Length 16;

Best Local Similarity 100.0%; Pred. No. 32;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 VAETPT 35

Db 6 vaetpt 11

RESULT 42

W88882

ID W88882 standard; Protein; 37 AA.

XX W88882;

XX 01-MAR-1999 (first entry)

XX Polypeptide fragment encoded by gene 100.

XX Human; secreted protein; fusion protein; gene therapy; protein therapy;

KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;

KW developmental abnormality; foetal deficiency; blood; allergy; renal;

KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;

KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;

KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;

KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;

KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

XX

```
OS Homo sapiens.
XX W09854963-A2.
XX PD 10-DEC-1998.
XX PF 04-JUN-1998; 98WO-US11422.
XX PR 18-DEC-1997; 97US-0070923.
PR 06-JUN-1997; 97US-0048877.
PR 06-JUN-1997; 97US-0048881.
PR 06-JUN-1997; 97US-0048884.
PR 06-JUN-1997; 97US-0048893.
PR 06-JUN-1997; 97US-0048896.
PR 06-JUN-1997; 97US-0048899.
PR 06-JUN-1997; 97US-0048915.
PR 06-JUN-1997; 97US-0048949.
PR 06-JUN-1997; 97US-0048964.
PR 06-JUN-1997; 97US-0048972.
PR 06-JUN-1997; 97US-0049020.
PR 06-JUN-1997; 97US-0049375.
PR 05-SEP-1997; 97US-0057628.
PR 05-SEP-1997; 97US-0057635.
PR 05-SEP-1997; 97US-0057644.
PR 05-SEP-1997; 97US-0057647.
PR 05-SEP-1997; 97US-0057650.
PR 05-SEP-1997; 97US-0057661.
PR 05-SEP-1997; 97US-0057667.
PR 05-SEP-1997; 97US-0057761.
PR 05-SEP-1997; 97US-0057764.
PR 05-SEP-1997; 97US-0057770.
PR 05-SEP-1997; 97US-0057775.
PR 06-JUN-1997; 97US-0048875.
PR 06-JUN-1997; 97US-0048878.
PR 06-JUN-1997; 97US-0048882.
PR 06-JUN-1997; 97US-0048885.
PR 06-JUN-1997; 97US-0048894.
PR 06-JUN-1997; 97US-0048897.
PR 06-JUN-1997; 97US-0048900.
PR 06-JUN-1997; 97US-0048916.
PR 06-JUN-1997; 97US-0048962.
PR 06-JUN-1997; 97US-0048970.
PR 06-JUN-1997; 97US-0048974.
PR 06-JUN-1997; 97US-0049373.
PR 05-SEP-1997; 97US-0057584.
PR 05-SEP-1997; 97US-0057629.
PR 05-SEP-1997; 97US-0057642.
PR 05-SEP-1997; 97US-0057645.
PR 05-SEP-1997; 97US-0057651.
PR 05-SEP-1997; 97US-0057662.
PR 05-SEP-1997; 97US-0057668.
PR 05-SEP-1997; 97US-0057762.
PR 05-SEP-1997; 97US-0057765.
PR 05-SEP-1997; 97US-0057771.
PR 05-SEP-1997; 97US-0057776.
PR 06-JUN-1997; 97US-0048876.
PR 06-JUN-1997; 97US-0048876.
PR 06-JUN-1997; 97US-0048880.
PR 06-JUN-1997; 97US-0048883.
PR 06-JUN-1997; 97US-0048892.
PR 06-JUN-1997; 97US-0048895.
PR 06-JUN-1997; 97US-0048898.
PR 06-JUN-1997; 97US-0048901.
PR 06-JUN-1997; 97US-0048917.
PR 06-JUN-1997; 97US-0048963.
PR 06-JUN-1997; 97US-0048971.
PR 06-JUN-1997; 97US-0049019.
PR 05-SEP-1997; 97US-0049374.
PR 05-SEP-1997; 97US-0057627.
PR 05-SEP-1997; 97US-0057634.
PR 05-SEP-1997; 97US-0057643.
PR 05-SEP-1997; 97US-0057646.

PR 05-SEP-1997; 97US-0057649.
PR 05-SEP-1997; 97US-0057654.
PR 05-SEP-1997; 97US-0057666.
PR 05-SEP-1997; 97US-0057760.
PR 05-SEP-1997; 97US-0057763.
PR 05-SEP-1997; 97US-0057769.
PR 05-SEP-1997; 97US-0057774.
PR 05-SEP-1997; 97US-0057777.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Brewer LA, Carter KC, Dillon PJ, Ebner R, Endress GA;
PI Fan P, Feng P, Ferrie AM, Fischer CL, Florence C;
PI Florence K, Greene JM, Hu J, Kyaw H, Lafleur DM;
PI Li Y, Moore PA, Ni J, Olsen HS, Rosen CA, Ruben SM;
PI Shi Y, Soppet DR, Wei Y, Young P, Yu G, Zeng Z;
XX WPI; 1999-059865/05.
DR N-ESDB; V84510.
XX New isolated human genes and the secreted polypeptides they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
XX Disclosure; Page 77: 772pp; English.
XX The invention relates to nucleic acid sequences (V84411 to V84633)
CC encoding human secreted proteins (W88534 to W88756). The secreted protein
CC gene sequences are deposited with the ATCC under deposit numbers ATCC
CC 97979, 97974, 97975, 97976, 97977, 209007, 209008, 209009, 209010,
CC 209011, 209080, 209081, 209082, 209083, 209084, 209085, 209511. Host
CC cells comprising recombinant vectors containing the nucleic acid
CC sequences are used for the recombinant production of the secreted
CC proteins. The polynucleotide and amino acid sequences are useful for are
CC useful for preventing, treating or ameliorating medical conditions e.g.
CC by protein or gene therapy. Pathological conditions can be also diagnosed
CC by determining the amount of the new polypeptides in a sample or by
CC determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the polynucleotides, based on
CC which tissues they are most highly expressed in, and include developing
CC products for the diagnosis or treatment of cancer, neurodegenerative
CC disorders, developmental abnormalities and foetal deficiencies, blood
CC disorders, tumours, leukemias, diseases of the immune system, autoimmune
CC diseases, hepatic and renal disease, lymphomas, inflammation, allergies,
CC ischemic shock, Alzheimer's and cognitive disorders, schizophrenia,
CC restenosis, prostate diseases, obesity, disorders involving osteoclasts
CC such as osteoporosis, arthritis or malignancies, diseases of testes, lung
CC or thymus, digestive/endocrine disorders, infections and AIDS. The
CC polypeptides are also useful for identifying their binding partners.
CC The present sequence represents a polypeptide fragment encoded by a
CC gene of the invention (see descriptor line for gene number).
XX Sequence 37 AA;
SQ
Query Match 3.5%; Score 6; DB 20; Length 37;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 17 ALPALL 22
DB 15 alpall 20
RESULT 43
Y41113
ID Y41113 standard; peptide; 40 AA.
XX Y41113;
XX 17-JAN-2000 (first entry)
XX TNFR/NGFR cysteine-rich domain of T129 polypeptide.
```

```

XX Tumor necrosis factor receptor; TNF; T129 protein; TANGO 129; human;
KW cellular process; immunological disorder; abnormal lymphoid development;
KW thymic development; T-cell mediated immune response; humoral B cell;
XX skeletal muscle disorder; drug screening.
XX Homo sapiens.
XX OS
XX WO9952924-A1.
XX PN
XX 21-OCT-1999.
XX PD
XX 08-APR-1999; 99WO-US07832.
XX PF
XX 09-APR-1998; 98US-0057951.
XX PR
XX (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.
XX PA
XX Holtzman D;
XX PI
XX WPI; 1999-620368/53.
XX DR
XX New isolated tumor necrosis factor receptor member used to develop
PT products for treating, e.g. immunological disorders or disorders of the
PT skeletal muscle -
PT
XX Example 3; Fig 2; 118pp; English.
XX PS
XX The invention provides an isolated human tumor necrosis factor (TNF)
CC receptor member, T129 (also referred as TANGO 129). The T129 polypeptide
CC can be expressed by standard recombinant methodology. The T219
CC polypeptides are useful as modulating agents in regulating a variety of
CC cellular processes. Agents or modulators which have a stimulatory or
CC inhibitory effect on T129 activity (e.g. T129 gene expression) as
CC identified by a screening assay can be administered to individuals to
CC treat (prophylactically or therapeutically) disorders, e.g. an
CC immunological disorder associated with aberrant T129 activity, disorders
CC associated with abnormal lymphoid and/or thymic development, T-cell
CC mediated immune response, T-cell dependent help for B cells, and abnormal
CC humoral B cell activity, and possibly disorders of the skeletal muscle.
CC The products can also be used for detection, diagnosis, drug screening
CC and production of transgenic animals. The present sequence represents
CC a TNFR/NGFR cysteine-rich domain of T129 polypeptide.
XX SQ
SQ Sequence 40 AA;

Query Match 3.5%; Score 6; DB 20; Length 40;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 52 CPPTGF 57
Db 1 cpptgf 6

RESULT 44
ID Y11899 standard; Protein; 42 AA.
XX
XX Y11899;
XX
XX 18-JUN-1999 (first entry)
XX
XX Human 5' EST secreted protein SEQ ID NO: 499.
XX
XX Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition.
XX
XX

```

```

OS Homo sapiens.
XX
XX WO9906550-A2.
XX PN
XX 11-FEB-1999.
XX PD
XX 31-JUL-1998; 98WO-IB01232.
XX PF
XX 01-AUG-1997; 97US-0905144.
XX PR
XX (GEST ) GENSET.
XX PA
XX Duclert A, Dumas Milne Edwards J, Lacroix B;
XX PI
XX WPI; 1999-153780/13.
XX DR
XX N-PSDB; X40621.
XX
XX New isolated prostate-derived nucleic acids - used to develop
PT products which may have cytokine, immune regulatory, haematopoiesis
PT regulating, anti-inflammatory or tumour inhibition activity
PT
XX Claim 34; Page 615; 675pp; English.
XX PS
XX X40438 to X40715 represent 5' expressed sequence tags (ESTs) for human
CC secreted proteins expressed in prostate, and encode the proteins given in
CC Y11716 to Y11993 respectively. The proteins given represent the signal
CC peptide and an N-terminal fragment of a secreted protein. The nucleic
CC acid sequences can be used for producing secreted human gene products.
CC They can also be used to develop products for diagnosis and therapy. The
CC proteins obtained may have cytokine activity, cell proliferation and
CC differentiation activity, haematopoiesis regulating activity, tissue
CC growth regulating activity, reproductive hormone regulating activity,
CC chemotactic/chemokinetic activity, haemostatic and thrombolytic activity,
CC receptor/ligand activity, anti-inflammatory activity, tumour inhibition
CC activity or other activities. The products can be used in forensic, gene
CC therapy and chromosome mapping procedures. The sequences can also be used
CC for obtaining corresponding promoter sequences. The nucleic acids
CC encoding the signal peptides can be used for directing extracellular
CC secretion of a polypeptide or the insertion of a polypeptide into a
CC membrane, or importing a polypeptide into a cell.
XX SQ
SQ Sequence 42 AA;

Query Match 3.5%; Score 6; DB 20; Length 42;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 LVLALP 19
Db 22 lvalp 27

RESULT 45
ID Y12205 standard; Protein; 43 AA.
XX
XX Y12205;
XX
XX 18-JUN-1999 (first entry)
XX
XX Human 5' EST secreted protein SEQ ID NO: 518.
XX
XX Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition.
XX
XX
XX Homo sapiens.
XX

```

```

PN WO906554-A2.
XX
XX
PD 11-FEB-1999.
XX
PF 31-JUL-1998; 98WO-IB01238.
XX
PR 01-AUG-1997; 97US-0905134.
XX
XX (GEST ) GENSET.
XX
XX
PI Duclert A, Dumas Milne Edwards J, Lacroix B;
XX
XX WPI; 1999-153784/13.
DR N-PSDB; X41038.
XX
XX New nucleic acids encoding human secreted proteins - obtained from
PT cDNA libraries prepared from kidney, fetal kidney, dystrophic
PT muscle, muscle and heart tissue
XX
XX Claim 34; Page 586; 622pp; English.
XX
XX X40826 to X41093 represent 5' expressed sequence tags (ESTs) for human
CC secreted proteins, and encode the proteins given in Y01602 and
CC Y11994 to Y12260, respectively. The proteins given represent the signal
CC peptide and an N-terminal fragment of a secreted protein. The nucleic
CC acid sequences can be used for producing secreted human gene products.
CC They can also be used to develop products for diagnosis and therapy.
CC The proteins obtained may have cytokine activity, cell
CC proliferation/differentiation activity, haematopoiesis regulating
CC activity, tissue growth regulating activity, reproductive hormone
CC regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, tumour inhibition activity or other activities. The products
CC can be used in forensic, gene therapy and chromosome mapping procedures.
CC The sequences can also be used for obtaining corresponding promoter
CC sequences. The nucleic acids encoding the signal peptide can be used
CC for directing extracellular secretion of a polypeptide or the insertion
CC of a polypeptide into a membrane, or importing a polypeptide into
XX a cell.
XX
XX Sequence 43 AA;
SQ
Query Match 3.5%; Score 6; DB 20; Length 43;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 19 PALLPV 24
Db 2 pallpv 7
|||||
RESULT 46
R07283
ID R07283 standard; peptide; 44 AA.
XX
XX R07283;
AC
XX
XX 31-JAN-1991 (first entry)
DT
XX
XX Smooth muscle myosin-1 immunogen for antibody prodn.
DE
XX
XX Smooth muscle myosin isoform 1; monoclonal antibody; immunogen;
KW diagnosis; arteriosclerosis.
XX
XX Homo sapiens.
OS
XX
XX WO9011520-A.
PN
XX
XX 04-OCT-1990.
PD
XX
XX 26-MAR-1990; 90WO-JP00398.
PF
XX
XX
Query Match 3.5%; Score 6; DB 11; Length 44;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 152 RSGGRR 157
Db 13 rsggrr 18
|||||
RESULT 47
Y27662
ID Y27662 standard; Protein; 50 AA.
XX
XX Y27662;
AC
XX
XX 30-JUL-1999 (first entry)
DT
XX
XX Human secreted protein encoded by gene No. 96.
DE
XX
XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
XX Homo sapiens.
OS
XX
XX WO9924836-A1.
PN
XX
XX 20-MAY-1999.
PD
XX
XX 04-NOV-1998; 98WO-US23435.
PF
XX
XX 17-NOV-1997; 97US-0066100.
PR
XX 07-NOV-1997; 97US-0064900.
PR
XX 07-NOV-1997; 97US-0064908.
PR
XX 07-NOV-1997; 97US-0064911.
PR
XX 07-NOV-1997; 97US-0064912.
PR
XX 07-NOV-1997; 97US-0064983.
PR
XX 07-NOV-1997; 97US-0064984.
PR
XX 07-NOV-1997; 97US-0064985.
PR
XX 07-NOV-1997; 97US-0064987.
PR

```

The oligopeptide is used as immunogen for the prodn. of monoclonal antibodies recognising isoform SM-1 of the heavy chain of smooth muscle myosin, pref. from heart or skeletal muscle, esp. human. The peptide contains the part which differs between isoforms SM1-3. The antibodies may be obtained by immunisation with the immunogen, followed by cell fusion to produce a hybridoma, cloning and culturing the chosen hybridoma clone. The Ab is a reagent for the histological staining of smooth muscle, and is useful in the diagnosis of arteriosclerosis, blood vessel disorders etc.

See also R07284-5.

Sequence 44 AA;

Query Match 3.5%; Score 6; DB 11; Length 44;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 RSGGRR 157
Db 13 rsggrr 18
|||||

RESULT 47
Y27662
ID Y27662 standard; Protein; 50 AA.
XX
XX Y27662;
AC
XX
XX 30-JUL-1999 (first entry)
DT
XX
XX Human secreted protein encoded by gene No. 96.
DE
XX
XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
XX Homo sapiens.
OS
XX
XX WO9924836-A1.
PN
XX
XX 20-MAY-1999.
PD
XX
XX 04-NOV-1998; 98WO-US23435.
PF
XX
XX 17-NOV-1997; 97US-0066100.
PR
XX 07-NOV-1997; 97US-0064900.
PR
XX 07-NOV-1997; 97US-0064908.
PR
XX 07-NOV-1997; 97US-0064911.
PR
XX 07-NOV-1997; 97US-0064912.
PR
XX 07-NOV-1997; 97US-0064983.
PR
XX 07-NOV-1997; 97US-0064984.
PR
XX 07-NOV-1997; 97US-0064985.
PR
XX 07-NOV-1997; 97US-0064987.
PR

PR 07-NOV-1997; 97US-0064988.
 PR 17-NOV-1997; 97US-0066090.
 PR 17-NOV-1997; 97US-0066094.
 PR 17-NOV-1997; 97US-0066095.
 PR 17-NOV-1997; 97US-0066089.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Carter KC, Ebner R, Endress GA, Peng P, Janat F;
 PI Kyaw H, Lafleur DW, Moore PA, Ni J, Olsen HS, Rosen CA;
 PI Ruben SM, Shi Y, Soppet DR, Wei Y;
 XX
 DR WPI; 1999-337740/28.
 DR N-PSDB; X85028.
 XX
 PT New human secreted proteins and coding sequences useful for treating
 PT disorders of the immune system and hyperproliferative disorders
 XX
 XX Claim 11; Page 402; 507pp; English.
 CC This sequence represents a secreted human protein encoded by the gene
 CC clone detailed in the descriptor line. The gene can be used to generate
 CC fusion proteins by linking to the gene to a human immunoglobulin Fc
 CC portion (e.g. X84924) for increasing the stability of the fused protein
 CC as compared to the human protein only.
 CC The invention relates to 125 novel genes and their fragments (nucleic
 CC acid sequences: X84933-X85057; amino acid sequences Y27567-Y27933) which
 CC are useful for preventing, treating or ameliorating medical conditions
 CC e.g. by protein or gene therapy. Also, pathological conditions can be
 CC diagnosed by determining the amount of the new polypeptides in a sample
 CC or by determining the presence of mutations in the new polynucleotides.
 CC Specific uses are described for each of the 125 polynucleotides, based on
 CC which tissues they are most highly expressed in (see X84933 for described
 CC uses).
 XX
 SQ Sequence 50 AA:
 Query Match 3.5%; Score 6; DB 20; Length 50;
 Best Local Similarity 100.0%; Pred. No. 86;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 9 LSLCL 14
 Db 25 LSLCL 30
 RESULT 48
 ID Y20790 standard; Protein; 53 AA.
 XX
 AC Y20790;
 XX
 XX 22-JUL-1999 (first entry)
 DE Human neurofilament-M mutant protein fragment 72.
 XX
 KW Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
 KW frameshift mutation; age-related disease; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN W09845322-A2.
 XX

PD 15-OCT-1998.
 XX
 PF 02-APR-1998; 98WO-1B00705.
 XX
 PR 10-APR-1997; 97US-0043163.
 XX
 PA (UYUT-) RIJKSUNIV UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.
 XX
 PI Burbach JPH, Grosveld FG, Van Leeuwen FW;
 DR WPI; 1998-609901/51.
 DR N-PSDB; X75759.
 XX
 PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA
 PS Disclosure; Figure 8; 258pp; English.
 XX
 CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.
 XX
 SQ Sequence 53 AA:
 Query Match 3.5%; Score 6; DB 19; Length 53;
 Best Local Similarity 100.0%; Pred. No. 90;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 152 RSGRR 157
 Db 41 rsgrr 46
 RESULT 49
 ID Y13123 standard; Protein; 57 AA.
 XX
 AC Y13123;
 XX
 DT 22-JUN-1999 (first entry)
 XX
 DE Human secreted protein encoded by 5' EST SEQ ID NO: 137.
 XX
 KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
 KW forensic; gene therapy; chromosome mapping; signal peptide;
 KW upstream regulatory sequence; cytokine activity; cell proliferation;
 KW differentiation; haematopoiesis regulation; tissue growth regulation;
 KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; anti-inflammatory; tumour inhibition.
 XX
 OS Homo sapiens.
 OS
 PN

```
PN WO906552-A2.
XX
PD 11-FEB-1999.
XX
PF 31-JUL-1998; 98WO-TB01236.
XX
PR 01-AUG-1997; 97US-0905223.
XX
PA (GEST ) GENSET.
XX
PI Duclert A, Dumas Milne Edwards J, Lacroix B;
XX
XX WPI; 1999-153782/13.
XX
DR N-PSDB; X51923.
XX
XX New isolated brain-derived nucleic acids - used to develop products
PT which may have cytokine, immune, regulatory, haematopoiesis
PT regulating, anti-inflammatory or tumour inhibition activity
XX
XX Claim 34; Page 519; 577pp; English.
XX
XX X51787 to X52019 represent 5' expressed sequence tags (ESTs) for human
CC secreted proteins, and encode the proteins given in Y12987 to Y13219,
CC respectively. The proteins given represent the signal peptide and an
CC N-terminal fragment of a secreted protein. The nucleic acid sequences
CC can be used for producing secreted human gene products. They can also
CC be used to develop products for diagnosis and therapy. The proteins
CC obtained may have cytokine activity, cell proliferation/differentiation
CC activity, haematopoiesis regulating activity, tissue growth regulating
CC activity, reproductive hormone regulating activity, chemotactic/
CC chemokinetic activity, haemostatic and thrombolytic activity, receptor/
CC ligand activity, anti-inflammatory activity, tumour inhibition activity
CC or other activities. The products can be used in forensic, gene therapy
CC and chromosome mapping procedures. The sequences can also be used for
CC obtaining corresponding promoter sequences. The nucleic acids encoding
CC the signal peptide can be used for directing extracellular secretion of
CC a polypeptide or the insertion of a polypeptide into a membrane, or
CC importing a polypeptide into a cell.
XX
XX Sequence 57 AA;
XX
XX Query Match 3.5%; Score 6; DB 20; Length 57;
XX Best Local Similarity 100.0%; Pred. No. 96;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 19 PALLPV 24
Db 25 pallpv 30
XX
XX RESULT 50
XX Y21183
XX ID Y21183 standard; Protein; 58 AA.
XX AC Y21183;
XX
XX 22-JUL-1999 (first entry)
XX
XX Human bcl2 proto-oncogene mutant protein fragment 31.
XX
XX Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
KW frameshift mutation; age-related disease; neurodegenerative disorder;
KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
KW neurofilament-F; presenilin 1; presenilin II; cellular tumour antigen;
KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
KW high mobility group protein-C; neuroendocrine specific protein A.
XX
XX Synthetic.
XX
OS Homo sapiens.
XX
XX WO9845322-A2.
XX
XX 15-OCT-1998.
PD
XX
XX 02-APR-1998; 98WO-IB00705.
PF
XX
XX 10-APR-1997; 97US-0043163.
PR
XX
XX (UYUT-) RIJKSUNIV UTRECHT.
PA
XX (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
PA
XX (UYRO-) UNIV ROTTERDAM ERASMUS.
XX
XX Burbach JPH, Grosveld FG, Van Leeuwen FW;
PI
XX WPI; 1998-609901/51.
XX
XX N-PSDB; X75766.
DR
XX
XX Diagnosing disease by detecting frameshift mutations in RNA or
PT corresponding protein mutations - used to diagnose cancer and
PT neurological diseases, particularly Alzheimer's disease, and also
PT for treatment and prevention with specific ribozymes or wild-type
PT RNA
XX
XX Disclosure; Figure 15; 258pp; English.
XX
XX This invention describes a novel method for the diagnosis of a disease
CC caused by or associated with, an RNA molecule that has a frameshift
CC mutation. The method is used to diagnose age-related diseases, especially
CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
CC and many others listed) or susceptibility to these disorders. The method
CC allows a definitive diagnosis of Alzheimer's disease in living patients,
CC at an early stage. It is based on the observation that disease may be
CC caused by mutations in RNA rather than DNA. The invention describes the
CC use of neuronal system RNA molecules, specifically proteins including
CC beta-amyloid precursor protein (beta-APP), the microtubule associated
CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
CC neurofilament-F, presenilin 1, presenilin II, glial fibrillary acidic
CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
CC protein-C (HMGP-C) and neuroendocrine specific protein A.
XX
XX Sequence 58 AA;
XX
XX Query Match 3.5%; Score 6; DB 19; Length 58;
XX Best Local Similarity 100.0%; Pred. No. 97;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 164 AGPSLA 169
Db 1 agpsla 6
XX
XX Search completed: January 30, 2001, 17:06:50
XX Job time: 1131 sec
```


THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 17:07:54 ; Search time 27.49 Seconds

(without alignments)
419.902 Million cell updates/sec

Title: US-09-518-931-4

Perfect score: 170

Sequence: 1 MRALEGPGLSLLCLVLPALP.....PRSGRGRCRGQVAGPSLAP 170

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 195891 seqs, 67900655 residues

Word size : 6

Total number of hits satisfying chosen parameters: 528

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

PIR_66:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	5.3	327	2 H83483	probable transmembr
2	8	4.7	440	2 T50912	hypothetical prote
3	8	4.7	530	2 B64905	sugar kinase homol
4	8	4.7	1172	1 TSH0P2	thrombospondin 2 p
5	5	7.4	2 S13515	retinoic acid rece	
6	7	4.1	89	2 S13517	retinoic acid rece
7	7	4.1	106	2 G82729	hypothetical prote
8	7	4.1	113	1 TMECEL	colicin E1 immunit
9	7	4.1	113	2 I64785	imm protein - Esch
10	7	4.1	113	2 S11532	colicin E1 immunit
11	7	4.1	168	2 E83573	hypothetical prote
12	7	4.1	274	2 C71685	hypothetical prote
13	7	4.1	292	2 JE0233	troponin-I - scall
14	7	4.1	292	2 T36914	probable integral
15	7	4.1	317	2 S51572	moca protein - Rhi
16	7	4.1	335	2 A72065	GTP binding protei
17	7	4.1	338	2 T26399	hypothetical prote
18	7	4.1	338	2 H72115	cationic amino aci
19	7	4.1	343	2 G81602	GNP1/DBG family pr
20	7	4.1	354	2 B55733	G protein-coupled
21	7	4.1	364	1 JE0292	fungal stress prot
22	7	4.1	393	2 E72549	probable acyl-CoA
23	7	4.1	407	2 T471665	bicyclomycin resis
24	7	4.1	414	2 T47281	hypothetical prote
25	7	4.1	432	2 T35527	hypothetical prote
26	7	4.1	433	2 T36609	probable secreted
27	7	4.1	434	2 S29612	pectate lyase (EC
28	7	4.1	453	2 H70333	conserved hypothet
29	7	4.1	455	2 B83213	probable MFS trans

hypothetical prote
vgJE protein - Esc
alpha-globulin typ
hypothetical prote
histidine kinase h
hypothetical prote
hypothetical prote
anthranilate synth
probable membrane
quinohemoprotein e
conserved hypothet
outer membrane ush
hypothetical prote
mitogen-activated
sodium channel mNa
nascent polypeptid
apolipoprotein B-1
hydrogenase (EC 1.
hypothetical prote
protein YBR126w-a
NuMA protein - hum
xylulokinase (EC 2
hypothetical prote
gp30 protein - Myc
embryonic abundant
hypothetical prote
ferradoxin 2(4Fe-4
probable DNA-bind
conserved hypothet
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
oxytocin / neuroph
oxytocin / neuroph
thrombospondin 2 -
hypothetical prote
probable membrane
hypothetical prote
hypothetical prote
cell adhesion prot
insulin-like pepti
insulin-like pepti
hypothetical prote
hypothetical prote
hypothetical 15.1K
RNA binding protei
RNA-binding protei
hypothetical prote
hypothetical prote
hypothetical prote
probable regulator
hypothetical prote
hypothetical prote
hypothetical prote
lactoylglutathione
conserved hypothet
gene B protein - p
hypothetical prote
hypothetical prote
hypothetical prote
conserved hypothet
conserved hypothet
hypothetical prote

103	6	3.5	172	1	QOBED7	HHLF4 protein - hu	176	2	C83280	probable transcript
104	6	3.5	172	2	E83301	hypothetical prote	177	2	T16201	hypothetical prote
105	6	3.5	174	2	T15176	hypothetical prote	178	1	J01874	Btl protein - toma
106	6	3.5	175	2	S77306	hypothetical prote	179	2	T26084	hypothetical prote
107	6	3.5	181	3	T41517	wp repeat protein	180	2	A60364	tropomyosin - mlg
108	6	3.5	186	2	C75095	hypothetical prote	181	2	E83234	hypothetical prote
109	6	3.5	189	2	T7809	hypothetical prote	182	2	JH0332	Iga (Fc) receptor,
110	6	3.5	190	2	G83377	hypothetical prote	183	2	S35022	nolt protein - Rhi
111	6	3.5	190	2	T35951	probable SIR2 fami	184	2	T40637	40s ribosomal prot
112	6	3.5	191	2	T35508	hypothetical prote	185	2	E70721	hypothetical prote
113	6	3.5	193	2	G70884	probable ISI602res	186	1	QOCVPY	Btl protein - pota
114	6	3.5	194	2	C72554	hypothetical prote	187	1	QOCVBV	Btl protein - squa
115	6	3.5	195	2	JC2258	substrate protein	188	2	JQ2305	BRI protein - pepp
116	6	3.5	195	2	E83298	hypothetical prote	189	2	S29384	cyclohexadienyl de
117	6	3.5	198	2	B71072	hypothetical prote	190	2	B64674	conserved hypotHet
118	6	3.5	202	2	H70909	probable resolvase	191	2	G71841	hypothetical prote
119	6	3.5	203	2	H70522	probable resolvase	192	2	E70725	probable linB prot
120	6	3.5	203	2	JC4645	cardiotrophin-1 -	193	2	T17321	hypothetical prote
121	6	3.5	203	2	B70986	probable isomerase	194	2	A58928	phosducin-like pro
122	6	3.5	206	1	A55169	pre-T-cell recepto	195	2	G83344	probable permease
123	6	3.5	206	2	T47433	hypothetical prote	196	2	H82964	oxidoreductase Rmd
124	6	3.5	206	2	D82173	conserved hypotHet	197	2	QREB08	oligopeptide trans
125	6	3.5	207	2	D82951	conserved hypotHet	198	2	E36283	oligopeptide trans
126	6	3.5	208	2	C75041	orotidine-5'-monop	199	2	E64184	oligopeptide trans
127	6	3.5	218	2	S74867	hypothetical prote	200	2	C82242	oligopeptide ABC t
128	6	3.5	222	2	S70878	galactosyl transfe	201	2	H82743	methionyl-tRNA for
129	6	3.5	223	2	C72733	hypothetical prote	202	2	T43947	N-acetyl-gamma-glu
130	6	3.5	225	2	E69896	hypothetical prote	203	2	CRH04	carbonate dehydrat
131	6	3.5	227	2	C75582	conserved hypotHet	204	2	G71131	probable ATP-bindi
132	6	3.5	228	2	S58986	cytochrome-c oxida	205	2	E83624	probable transcript
133	6	3.5	228	2	A72740	hypothetical prote	206	2	B83487	hypothetical prote
134	6	3.5	231	2	B49934	phosphoglycolate p	207	2	T34838	probable transera
135	6	3.5	231	2	D49934	phosphoglycolate p	208	2	D72700	hypothetical prote
136	6	3.5	231	2	D81441	probable integral	209	2	I37371	ER calcium-binding
137	6	3.5	232	2	JQ1199	replication protei	210	2	C64852	probable pseudouri
138	6	3.5	233	2	S33946	hexon-associated p	211	2	JC5402	vitamin D receptor
139	6	3.5	233	2	S15954	repB protein - Lac	212	2	C82170	peptide ABC transp
140	6	3.5	236	2	T36188	hypothetical prote	213	2	T02987	myb-related protei
141	6	3.5	238	2	F71723	ribosomal protein	214	2	T08462	hypothetical prote
142	6	3.5	238	2	F81151	hypothetical prote	215	2	S15201	hydrogenase expres
143	6	3.5	238	2	S12259	hypothetical prote	216	2	D70125	Lambda CII stabili
144	6	3.5	239	2	G02630	FcalphaRb - human	217	2	T08715	hypothetical prote
145	6	3.5	239	2	G83143	hypothetical prote	218	1	IMBKBB	immunogenic protei
146	6	3.5	240	2	F70443	6-carboxyhexanoate	219	2	I65768	smooth muscle myos
147	6	3.5	241	2	D71120	hypothetical prote	220	2	F75473	conserved hypotHet
148	6	3.5	241	2	S53522	crystallin beta B1	221	2	G83098	hypothetical prote
149	6	3.5	242	2	H72512	hypothetical prote	222	2	C75338	conserved hypotHet
150	6	3.5	244	2	A41021	NAD+-asparagine A	223	2	T26349	conserved hypotHet
151	6	3.5	247	2	A48982	glutathione transf	224	2	G69530	hypothetical prote
152	6	3.5	248	2	T14957	hypothetical prote	225	2	T15711	hypothetical prote
153	6	3.5	251	2	B70765	probable cobM - My	226	2	S38147	spas protein - Sal
154	6	3.5	251	2	A38912	NAD+-asparagine A	227	2	S37356	hypothetical prote
155	6	3.5	253	2	T27951	hypothetical prote	228	2	S38355	hypothetical prote
156	6	3.5	253	2	H70780	hypothetical prote	229	2	T45956	hypothetical prote
157	6	3.5	253	2	F72388	prephenate dehydro	230	2	T70859	hypothetical prote
158	6	3.5	256	2	H71537	hypothetical prote	231	2	B83152	hypothetical prote
159	6	3.5	257	1	S40985	H+-transporting At	232	2	C69005	succinyl-CoA synth
160	6	3.5	257	2	S71335	alcohol dehydrogen	233	2	C69005	hepatic transcript
161	6	3.5	258	2	B61228	collagen alpha 1(I	234	2	A34894	2-aminoethylphosph
162	6	3.5	260	2	T41687	conserved hypotHet	235	2	H82437	G protein-coupled
163	6	3.5	262	2	S72890	hypothetical prote	236	2	I52315	hypothetical prote
164	6	3.5	265	2	T36391	hypothetical prote	237	2	E72595	hypothetical prote
165	6	3.5	265	2	T03523	hypothetical prote	238	2	H81783	hypothetical prote
166	6	3.5	265	2	S39874	hypothetical prote	239	2	B83547	citrate synthase 2
167	6	3.5	268	2	B39429	CAMP response elem	240	2	OKKW1R	protein kinase (EC
168	6	3.5	268	2	B37279	enhancer-binding p	241	2	T16701	hypothetical prote
169	6	3.5	268	2	C37280	C/EBP-related prot	242	2	B71652	iron-sulfur cofact
170	6	3.5	268	2	S24377	hypothetical prote	243	2	S70841	hypothetical trans
171	6	3.5	269	2	T35676	cobalamin synthase	244	2	H70745	probable udp-gluc
172	6	3.5	271	2	T36685	probable membrane	245	2	B81206	hypothetical prote
173	6	3.5	273	2	H70849	hypothetical prote	246	2	S39007	ubiquinol--cytochr
174	6	3.5	276	1	JC4005	glutamate racemase	247	2	B64473	hypothetical prote
175	6	3.5	276	2	H75588	conserved hypotHet	248	2	JC4378	metalloproteinase

249	6	3.5	382	2	D58930	ubiquinol--cytochr	322	6	3.5	473	2	S18315	ribulose-bisphosph
250	6	3.5	383	2	F70752	hypothetical prote	323	6	3.5	476	2	H82177	conserved hypoteth
251	6	3.5	384	2	T04427	hypothetical prote	324	6	3.5	480	2	H70854	hypothetical prote
252	6	3.5	385	2	F82030	probable GTP-bindi	325	6	3.5	482	2	T15829	hypothetical prote
253	6	3.5	386	2	T19201	hypothetical prote	326	6	3.5	483	2	S52974	regulatory protein
254	6	3.5	387	2	A82284	conserved hypoteth	327	6	3.5	484	2	S60943	RUD3 protein - yea
255	6	3.5	388	2	S37098	naringenin-chalcon	328	6	3.5	485	2	T35571	hypothetical prote
256	6	3.5	389	2	A64966	serine-type D-Ala-	329	6	3.5	488	2	B81831	probable integral
257	6	3.5	390	2	F0LJGA	gag polyprotein -	330	6	3.5	488	2	A81194	C4-dicarboxylate t
258	6	3.5	391	2	S04205	protein-tyrosine k	331	6	3.5	490	2	T26983	hypothetical prote
259	6	3.5	392	2	S29356	gag protein - bovi	332	6	3.5	491	2	E75486	hypothetical prote
260	6	3.5	393	2	S44277	flavohe moglobin hm	333	6	3.5	495	2	S60761	alpha-2.8 polysial
261	6	3.5	394	1	S15992	flavohe moglobin hm	334	6	3.5	498	2	A48203	interleukin-14 pre
262	6	3.5	395	2	D72687	probable mu-crysta	335	6	3.5	500	2	F83418	sodium/proton anti
263	6	3.5	396	2	D72687	GTP cyclohydrolase	336	6	3.5	501	2	B75462	hypothetical prote
264	6	3.5	397	1	S45545	hypothetical prote	337	6	3.5	502	2	T05246	cytochrome P450 83
265	6	3.5	400	2	H83536	beta-ketoacyl Co	338	6	3.5	503	2	A23547	keratin, type II c
266	6	3.5	401	2	D83618	hypothetical prote	339	6	3.5	503	2	S38005	hypothetical prote
267	6	3.5	405	2	S55606	protein C inhibito	340	6	3.5	506	1	WQEC2S	phosphotransferase
268	6	3.5	406	2	A39339	hypothetical prote	341	6	3.5	506	2	T18871	hypothetical prote
269	6	3.5	407	2	T24492	hypothetical 44.9K	342	6	3.5	510	2	F82523	hypothetical prote
270	6	3.5	410	2	S77661	hypothetical prote	343	6	3.5	513	2	C81736	60 kDa chaperonin,
271	6	3.5	410	2	E70744	probable senx3 pro	344	6	3.5	514	2	B83360	hypothetical prote
272	6	3.5	412	2	S51760	ferredoxin reducta	345	6	3.5	515	2	G75267	ABC transporter, p
273	6	3.5	412	2	C83297	carboxypeptidase G	346	6	3.5	516	1	FWCNBB	beta-globulin B pr
274	6	3.5	413	2	A70795	hypothetical prote	347	6	3.5	520	2	T04550	hypothetical prote
275	6	3.5	413	2	T49165	hypothetical prote	348	6	3.5	521	2	T15322	hypothetical prote
276	6	3.5	416	2	G83108	probable type II s	349	6	3.5	523	2	S47053	hypothetical prote
277	6	3.5	416	2	T35282	probable two compo	350	6	3.5	524	2	I51268	cytochrome P450 ar
278	6	3.5	417	2	F71238	hypothetical prote	351	6	3.5	524	2	T46310	hypothetical prote
279	6	3.5	417	2	T29328	hypothetical prote	352	6	3.5	528	2	D70968	hypothetical prote
280	6	3.5	418	2	T47818	hypothetical prote	353	6	3.5	528	2	T34417	delayed rectifier
281	6	3.5	419	2	A37926	fumarylacetoacetas	354	6	3.5	529	2	C75453	GMC oxidoreductase
282	6	3.5	420	2	T43150	probable phosphori	355	6	3.5	531	2	S28260	achacin precursor
283	6	3.5	421	2	T30756	hypothetical prote	356	6	3.5	532	1	A29849	intercellular adhe
284	6	3.5	426	2	A69394	amine dehydrogenas	357	6	3.5	533	1	VHBPBL	minor capsid prote
285	6	3.5	430	2	A69778	metabolite transpo	358	6	3.5	536	2	E75580	adenine deaminase
286	6	3.5	431	2	F71238	hypothetical prote	359	6	3.5	539	2	T03469	probable hydrogena
287	6	3.5	434	1	S61999	hypothetical prote	360	6	3.5	539	2	I49065	lymphoid-restrict
288	6	3.5	435	2	H75006	hypothetical prote	361	6	3.5	540	2	T50617	hypothetical prote
289	6	3.5	436	2	T46107	methyl-accepting c	362	6	3.5	540	2	B47417	insulin receptor-r
290	6	3.5	437	2	T48035	hypothetical prote	363	6	3.5	546	2	S52053	cytochrome-c oxida
291	6	3.5	439	2	G70693	probable dinf prot	364	6	3.5	547	1	S51475	cytochrome P450 cy
292	6	3.5	440	2	H69989	lipoprotein homolo	365	6	3.5	547	2	G83116	pyochelin biosynth
293	6	3.5	441	2	H36789	hypothetical prote	366	6	3.5	555	2	JC4711	epoxide hydrolase
294	6	3.5	443	2	C70818	probable ABC trans	367	6	3.5	556	2	S49647	phosphoglycerate m
295	6	3.5	443	2	B83221	metalloproteinase	368	6	3.5	557	2	T47665	beta-N-acetylhexos
296	6	3.5	445	2	B82954	probable dihydrooru	369	6	3.5	559	2	S60473	phosphoglycerate m
297	6	3.5	446	2	G70510	probable oxidoredu	370	6	3.5	559	2	H70804	hypothetical prote
298	6	3.5	448	2	S12526	metalloproteinase	371	6	3.5	570	2	A82135	2-succinyl-6-hydro
299	6	3.5	449	2	S48173	variable surface g	372	6	3.5	574	2	T14162	hABC transport pro
300	6	3.5	450	2	B83530	probable RNA methy	373	6	3.5	581	2	E69322	dolichol-P-glucose
301	6	3.5	451	2	H82044	C4-dicarboxylate t	374	6	3.5	586	2	S64779	probable membrane
302	6	3.5	458	2	A75386	conserved hypoteth	375	6	3.5	587	2	A31776	lactose permease -
303	6	3.5	459	2	T29443	probable two-compo	376	6	3.5	588	1	FWCNAB	alpha-globulin B p
304	6	3.5	460	1	A34180	sulfite oxidase (E	377	6	3.5	591	2	S77707	phosphoribosylam
305	6	3.5	461	2	A46394	suppressor protein	378	6	3.5	592	1	IKBECA	colicin A - Citrob
306	6	3.5	461	2	G83502	hypothetical prote	379	6	3.5	592	2	S54489	phosphoribosylam
307	6	3.5	461	2	T47782	hypothetical prote	380	6	3.5	593	2	T24379	hypothetical prote
308	6	3.5	462	2	F83510	flagellar hook pro	381	6	3.5	593	2	B71323	conserved hypoteth
309	6	3.5	462	2	B83123	probable endoprote	382	6	3.5	594	2	S33561	ref(2)p protein -
310	6	3.5	463	2	A40013	phosphomannomutase	383	6	3.5	595	2	T06412	probable chaperoni
311	6	3.5	463	2	H82979	phosphomannomutase	384	6	3.5	605	2	A81278	hypothetical prote
312	6	3.5	465	2	JC4793	dihydrolipoamide d	385	6	3.5	607	2	B81278	hypothetical prote
313	6	3.5	465	2	A82211	MutI/nudix family	386	6	3.5	614	2	S59517	serum albumin prec
314	6	3.5	467	2	B65020	hypothetical prote	387	6	3.5	614	2	S27962	modulator recognit
315	6	3.5	468	2	H82204	probable multidrug	388	6	3.5	617	2	I56530	gene VGF protein -
316	6	3.5	469	2	D72722	probable MRSA prot	389	6	3.5	617	2	A39748	nerve growth facto
317	6	3.5	469	2	G65058	hypothetical prote	390	6	3.5	627	1	A43300	squalene-hopene cy
318	6	3.5	470	2	T22785	hypothetical prote	391	6	3.5	627	2	B71709	hypothetical prote
319	6	3.5	472	2	B26430	L-arabinose isomer	392	6	3.5	629	2	E70589	protoporphyrin IX
320	6	3.5	472	2	S47089	arabinose-proton s	393	6	3.5	637	2	S35221	globulin Begl prec
321	6	3.5	473	1	RKBLIT	ribulose-bisphosph	394	6	3.5	646	2	G70032	ABC transporter (p

395 646 2 B70396
396 647 2 G72346
397 654 2 S76870
398 656 2 D82388
399 663 1 TVMVR
400 668 1 A42908
401 671 2 A55023
402 674 2 J24258
403 676 2 E82954
404 684 2 I39595
405 684 2 E75434
406 687 2 T29148
407 690 2 S35211
408 690 2 B24099
409 695 2 S21433
410 718 2 T05840
411 730 1 I52580
412 730 2 J21456
413 732 2 I60729
414 732 2 S35631
415 733 2 T35429
416 733 2 T27289
417 738 2 H82807
418 740 2 A83271
419 741 2 T06017
420 746 2 T04188
421 747 2 T04188
422 749 2 A75560
423 751 2 T46517
424 755 2 T48553
425 755 2 A44315
426 756 2 T04187
427 756 2 T04186
428 769 2 T05839
429 771 2 B70564
430 772 2 T05838
431 773 2 I59351
432 775 2 A46627
433 775 2 T04189
434 776 2 T29064
435 777 2 T41075
436 783 2 T35389
437 788 2 G82764
438 798 2 T50514
439 803 2 T04190
440 809 2 S16266
441 818 2 B83537
442 826 2 T01494
443 830 2 T10545
444 833 2 JQ1866
445 835 2 S75842
446 845 2 E72652
447 845 2 S75644
448 848 2 C65083
449 854 1 TNBEA1
450 871 1 T3533
451 886 2 S54355
452 892 2 A56213
453 900 2 A42024
454 909 2 A73337
455 911 1 B3H0
456 915 2 T26695
457 915 2 T09575
458 919 2 JC5934
459 932 2 JC5953
460 933 2 S17233
461 942 2 S75598
462 944 2 S66870
463 946 2 I38100
464 946 2 T00024
465 956 1 A46016
466 956 2 A57121
467 956 2 B83200

468 3.5 971 2 T24866
469 3.5 975 1 A31497
470 3.5 977 2 G72526
471 3.5 1011 2 S75806
472 3.5 1016 2 T31343
473 3.5 1018 2 T15297
474 3.5 1018 2 T22318
475 3.5 1024 2 T16491
476 3.5 1044 2 T43800
477 3.5 1049 1 IJHUG1
478 3.5 1079 2 T38913
479 3.5 1091 2 T35822
480 3.5 1097 2 T31504
481 3.5 1110 2 T29327
482 3.5 1111 2 T01239
483 3.5 1134 2 G9269
484 3.5 1155 2 E82144
485 3.5 1165 2 D72496
486 3.5 1178 1 A39804
487 3.5 1197 1 A48350
488 3.5 1203 2 T04294
489 3.5 1222 2 T14805
490 3.5 1236 1 VHWWE
491 3.5 1239 1 VHWVEV
492 3.5 1241 2 S26373
493 3.5 1242 2 S23350
494 3.5 1242 2 A56605
495 3.5 1275 2 D38164
496 3.5 1287 2 T22235
497 3.5 1288 2 T37528
498 3.5 1294 2 T48349
499 3.5 1300 2 A36502
500 3.5 1300 2 T18364
501 3.5 1309 2 F82207
502 3.5 1315 2 T05300
503 3.5 1344 2 S47412
504 3.5 1360 2 T06699
505 3.5 1443 2 S05979
506 3.5 1513 2 A54895
507 3.5 1559 2 T07757
508 3.5 1607 2 T43212
509 3.5 1900 2 S45530
510 3.5 1905 2 T18267
511 3.5 1972 1 A41604
512 3.5 1977 2 S54771
513 3.5 1979 1 S03166
514 3.5 2082 2 T37056
515 3.5 2182 1 GNNYB1
516 3.5 2201 2 A54774
517 3.5 2224 1 KFHU5
518 3.5 2285 1 G02434
519 3.5 2297 2 T34918
520 3.5 2301 1 GNNYTN
521 3.5 2303 1 GNNYTM
522 3.5 2647 2 A37098
523 3.5 3033 1 GNNVJ8
524 3.5 3707 2 S18252
525 3.5 4391 2 A38096
526 3.5 4861 2 S71752
527 3.5 5376 2 T42215
528 3.5 5376 2 T42215

hypothetical prote
kinesin heavy chai
hypothetical prote
hemolysin secretio
proline dehydrogen
hypothetical prote
hypothetical prote
hypothetical prote
protein kinase bub
desmoglein 1 precu
translation initia
probable regulator
hypothetical prote
hypothetical prote
hypothetical prote
molybdopterin-bind
transcriptional repa
hypothetical prote
thrombospondin pre
DNA-binding protei
hypothetical prote
hypothetical prote
structural polypro
structural polypro
genome polyprotein
structural polypro
cobN protein - Pse
hypothetical prote
probable snf2 faml
EIN2 protein - Ara
insulin receptor-r
ro-3 protein - Neu
ATP-dependent heli
hypothetical prote
gene P2 protein -
zinc finger protei
steroid hormone re
mucin 2, intestina
probable DNA (cyto
insulin-like growt
probable 1-phospha
multidrug resistan
myosin heavy chain
sodium channel alp
myosin heavy chain
probable multi-dom
genome polyprotein
ATP binding casset
coagulation factor
DNA-directed DNA p
polyketide synthas
genome polyprotein
gelation factor AB
genome polyprotein
heparan sulfate pr
perlecan precursor
giant protein p619
zonadhesin - mouse

ALIGNMENTS

RESULT 1
H83483
Probable transmembrane sensor PA1301 [imported] - Pseudomonas aeruginosa (strain PAO1
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 15-Sep-2000
C:Accession: H83483
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
 ; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
 A:Reference number: A82950
 A:Accession: H83483
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-327 <STO>
 A:Cross-references: GB:AE004559; GB:AE004091; NID:g9947228; PIDN:ARG04690.1; GSPDB:GN001
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA1301

Query Match 5.3%; Score 9; DB 2; Length 327;
 Best Local Similarity 100.0%; Pred. No. 0.4; Mismatches 0; Indels 0; Gaps 0;
 Matches 9; Conservative 0;

Qy 16 LALPALLPV 24
 |||||
 Db 304 LALPALLPV 312

RESULT 2
 T50912
 hypothetical protein ORF440 [imported] - Rubrivivax gelatinosus
 C:Species: Rubrivivax gelatinosus
 C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
 C:Accession: T50912
 R:Nagashima, K.V.; Igarashi, N.; Harada, J.; Nagashima, S.; Matsuura, K.; Shimada, K.
 submitted to the EMBL Data Library, November 1999
 A:Description: Determination of Nucleotide Sequences of Rubrivivax gelatinosus Photosynt
 A:Reference number: 225270
 A:Accession: T50912
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-440 <NAG>
 A:Cross-references: EMBL:AB034704; PIDN:BAA94065.1
 A:Experimental source: strain IL144
 C:Genetics:
 A:Note: ORF440

Query Match 4.7%; Score 8; DB 2; Length 440;
 Best Local Similarity 100.0%; Pred. No. 4.8; Mismatches 0; Indels 0; Gaps 0;
 Matches 8; Conservative 0;

Qy 16 LALPALLPV 23
 |||||
 Db 42 LALPALLPV 49

RESULT 3
 B64905
 sugar kinase homolog ydev - *Escherichia coli*
 C:Species: *Escherichia coli*
 C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 17-Mar-2000
 C:Accession: B64905
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
 .A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of *Escherichia coli* K-12.
 A:Reference number: A64720; MUID:97428617
 A:Accession: B64905
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-530 <BLAT>
 A:Cross-references: GB:AE000248; GB:U00096; NID:g1787783; PIDN:AAC74584.1; PID:g1787789;
 A:Experimental source: strain K-12, substrain MG1655
 C:Genetics:
 A:Gene: ydev
 C:Superfamily: xylulokinase

Query Match 4.7%; Score 8; DB 2; Length 530;
 Best Local Similarity 100.0%; Pred. No. 5.5; Mismatches 0; Indels 0; Gaps 0;
 Matches 8; Conservative 0;

Qy 41 AETGERLV 48
 |||||
 Db 475 AETGERLV 482

RESULT 4
 TSHUP2
 thrombospondin 2 precursor - human
 C:Species: Homo sapiens (man)
 C:Date: 19-May-1995 #sequence_revision 03-Aug-1995 #text_change 13-Aug-1999
 C:Accession: A47379; A42173
 R:LaBell, T.L.; Byers, P.H.
 Genomics 17, 225-229, 1993
 A:Title: Sequence and characterization of the complete human thrombospondin 2 cDNA: p
 A:Reference number: A47379; MUID:94010892
 A:Accession: A47379
 A:Molecule type: mRNA
 A:Residues: 1-1172 <LAB>
 A:Cross-references: GB:L12350; NID:g307505; PIDN:AAA03703.1; PID:g307506
 R:LaBell, T.L.; Milewicz, D.J.; Distech, C.M.; Byers, P.H.
 Genomics 12, 421-429, 1992
 A:Title: Thrombospondin II: partial cDNA sequence, chromosome location, and expressio
 A:Reference number: A42173; MUID:92217961
 A:Accession: A42173
 A:Molecule type: mRNA
 A:Residues: 560-1172 <LA2>
 A:Cross-references: GB:M81339
 A:Experimental source: fibroblast
 A:Note: sequence extracted from NCBI backbone (NCBIN:95091, NCBI:95096)
 C:Genetics:
 A:Gene: GDB:THBS2; TSP2
 A:Cross-references: GDB:128789; OMIM:188061
 A:Map position: 6q27-6q27
 C:Complex: homotrimer, disulfide linked
 C:Function:
 A:Description: participates in cell migration and adhesion, and in platelet aggregati
 C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology;
 C:Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; tri
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-1172/Product: thrombospondin 2 #status predicted <MAT>
 F:319-377/Domain: von Willebrand factor type C repeat homology <VWC>
 F:380-431/Domain: thrombospondin type 1 repeat homology <THR1>
 F:436-492/Domain: thrombospondin type 1 repeat homology <THR2>
 F:493-549/Domain: thrombospondin type 1 repeat homology <THR3>
 F:553-588/Domain: EGF homology <EGF1>
 F:652-691/Domain: EGF homology <EGF>
 F:928-930/Region: cell attachment (R-G-D) motif
 F:151,316,330,457,584,710,1069/Binding site: carbohydate (Asn) (covalent) #status pr
 F:167-226/Disulfide bonds: #status predicted
 F:266,270/Disulfide bonds: interchain #status predicted
 F:612/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match 4.7%; Score 8; DB 1; Length 1172;
 Best Local Similarity 100.0%; Pred. No. 10; Mismatches 0; Indels 0; Gaps 0;
 Matches 8; Conservative 0;

Qy 3 ALEGPGLS 10
 |||||
 Db 102 ALEGPGLS 109

RESULT 5
 S13515
 retinoic acid receptor beta-1 - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Aug-1999

C:Accession: S13515
R:Zelent, A.; Mendelssohn, C.; Kastner, P.; Krust, A.; Garnier, J.M.; Ruffenach, F.; Lerch
EMBO J. 10, 71-81, 1991
A:Title: Differentially expressed isoforms of the mouse retinoic acid receptor beta are
C:Superfamily: retinoic acid receptor alpha; erba transforming protein homology
C:Keywords: DNA binding; nucleus; transcription regulation; zinc finger

Query Match 4.1%; Score 7; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 12; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0

QY 23 PVPVRG 29
Db 9 PVPVRG 15
|||||

RESULT 6
S13517
retinoic acid receptor beta-3 - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Aug-1999
C:Accession: S13517
R:Zelent, A.; Mendelssohn, C.; Kastner, P.; Krust, A.; Garnier, J.M.; Ruffenach, F.; Lerch
EMBO J. 10, 71-81, 1991
A:Title: Differentially expressed isoforms of the mouse retinoic acid receptor beta are
C:Superfamily: retinoic acid receptor alpha; erba transforming protein homology
C:Keywords: DNA binding; nucleus; transcription regulation; zinc finger

Query Match 4.1%; Score 7; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 14; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0

QY 23 PVPVRG 29
Db 9 PVPVRG 15
|||||

RESULT 7
G82729
hypothetical protein XF1055 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: G82729
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20355717
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: G82729
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-106 <STM>
A:Cross-references: GB:AE003942; GB:AE003849; NID:g9105990; PIDN:AAF83865.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Terro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, B

A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF1055

Query Match 4.1%; Score 7; DB 2; Length 106;
Best Local Similarity 100.0%; Pred. No. 16; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0

QY 39 RDAETGE 45
Db 87 RDAETGE 93
|||||

RESULT 8
IMECEL
colicin E1 immunity protein - Escherichia coli plasmid colicin E1
C:Species: Escherichia coli
C:Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 24-Sep-1999
C:Accession: A94079; B93118; I54985; A03513; A24685
R:Waleh, N.S.; Johnson, P.H.
Proc. Natl. Acad. Sci. U.S.A. 82, 8389-8393, 1985
A:Title: Structural and functional organization of the colicin E1 operon.
A:Reference number: A94079; MUID:86094231
A:Accession: A94079
A:Molecule type: DNA
A:Residues: 1-113 <WAL>
R:Oka, A.; Nomura, N.; Morita, M.; Sugisaki, H.; Sugimoto, K.; Takanami, M.
Mol. Gen. Genet. 172, 151-159, 1979
A:Title: Nucleotide sequence of small ColE1 derivatives: structure of the regions ess
A:Reference number: A93118; MUID:80010893
A:Accession: B93118
A:Molecule type: DNA
A:Residues: 1-113 <OKA>
A:Cross-references: GB:V00268; NID:g41124; PIDN:CAA23529.1; PID:g41125
A:Note: Plasmid pAO3 is a small col E1 derivative
R:Jilk, R.A.; Makris, J.C.; Borchardt, L.; Reznikoff, W.S.
J. Bacteriol. 175, 1264-1271, 1993
A:Title: Implications of Tn3-associated adjacent deletions.
A:Reference number: I54985; MUID:93186691
A:Accession: I54985
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-113 <RES>
A:Cross-references: GB:S56312; NID:g298343
C:Genetics:
A:Gene: imm
A:Map position: 75-100/0-1
A:Genome: plasmid
C:Superfamily: colicin E1 immunity protein

Query Match 4.1%; Score 7; DB 1; Length 113;
Best Local Similarity 100.0%; Pred. No. 17; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0

QY 11 LLCLVLA 17
Db 92 LLCLVLA 98
|||||

RESULT 9
I64785
imm protein - Escherichia coli plasmid R100
C:Species: Escherichia coli
C:Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 26-Aug-1999
C:Accession: I64785

R:Ohtsubo, H.; Ryder, T.B.; Maeda, Y.; Armstrong, K.A.; Ohtsubo, E.
Adv. Biophys. 21, 115-133, 1986

A:Title: DNA replication of the resistance plasmid R100 and its control.

A:Reference number: 151821; MUID:86319522

A:Accession: I64785

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-113 <RES>
A:Cross-references: GB:M26840; NID:g151770; PIDN:AAA26071.1; PID:g151777

C:Genetics:

A:Genome: plasmid

C:Superfamily: colicin E1 immunity protein

Query Match 4.1%; Score 7; DB 2; Length 113;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 LCLVLVA 17

Db 92 LCLVLVA 98

RESULT 10

S11532

colicin E1 immunity protein - Shigella sonnei plasmid pKY-1

C:Species: Shigella sonnei

C:Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 26-Aug-1999

C:Accession: S11532

R:Higashi, M.; Hata, M.; Hase, T.; Yamaguchi, K.; Masamune, Y.

J. Gen. Appl. Microbiol. 32, 433-442, 1986

A:Title: The nucleotide sequence of cea and the region of origin of plasmid pKY-1.

A:Reference number: S06218

A:Accession: S11532

A:Molecule type: DNA

A:Residues: 1-113 <JGE>

A:Cross-references: EMBL:M37218; NID:g149346; PIDN:AAA98157.1; PID:g149348

C:Genetics:

A:Gene: imm

A:Genome: plasmid

C:Superfamily: colicin E1 immunity protein

Query Match 4.1%; Score 7; DB 2; Length 113;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 LCLVLVA 17

Db 93 LCLVLVA 99

RESULT 11

E83573

hypothetical protein PA0585 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 15-Sep-2000

C:Accession: E83573

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A:Reference number: A82950

A:Accession: E83573

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-168 <STO>

A:Cross-references: GB:AF004494; GB:AE004091; NID:g9946446; PIDN:AAG03974.1; GSPDB:GN001

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA0585

Query Match 4.1%; Score 7; DB 2; Length 168;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 LCLVLVA 17

Db 8 LCLVLVA 14

RESULT 12

C71685

hypothetical protein RP296 - Rickettsia prowazekii

C:Species: Rickettsia prowazekii

C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 21-Jul-2000

C:Accession: C71685

R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark

Nature 396, 133-140, 1998

A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.

A:Reference number: A71630; MUID:99039499

A:Accession: C71685

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-274 <AND>

A:Cross-references: GB:AJ235271; GB:AJ235269; NID:g3868717; PIDN:CAA14757.1; PID:g386

A:Experimental source: strain Madrid E

C:Genetics:

A:Gene: RP296

Query Match 4.1%; Score 7; DB 2; Length 274;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 LSLCLLV 15

Db 9 LSLCLLV 15

RESULT 13

JE0233

troponin-I - scallop (Chlamys nipponensis)

C:Species: Chlamys nipponensis (Japanese scallop)

C:Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 17-Mar-2000

C:Accession: JE0233

R:Tanaka, H.; Ojima, T.; Nishita, K.

J. Biochem. 124, 304-310, 1998

A:Title: Amino acid sequence of troponin-I from Akazara scallop striated adductor mus

A:Reference number: JE0233; MUID:98351986

A:Accession: JE0233

A:Molecule type: protein

A:Residues: 1-292 <TAB>

C:Superfamily: troponin T

Query Match 4.1%; Score 7; DB 2; Length 292;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 97 EREEAR 103

Db 121 EREEAR 127

RESULT 14

T36914

probable integral membrane sugar transport protein - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000

C:Accession: T36914

R:Seeger, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A

submitted to the EMBL Data Library, July 1999

A:Reference number: 221574

A:Accession: T36914
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-292 <SEE>
A:Cross-references: EMBL:AL096743; PIDN:CAB46401.1; GSPDB:GN00070; SCOEDB:SCI7.18
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SCI7.18
C:Superfamily: maltose transport protein malG

Query Match 4.1%; Score 7; DB 2; Length 292;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 LVLALPA 20
|||||
DB 103 LVLALPA 109

RESULT 15
S51572
mocA protein - Rhizobium meliloti
C:Species: Rhizobium meliloti
C:Date: 06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 08-Oct-1999
C:Accession: S51572; S43167
R:Rossbach, S.; Kulp, D.A.; Rossbach, U.; de Bruijn, F.J.
Mol. Gen. Genet. 245, 11-24, 1994
A:Title: Molecular and genetic characterization of the rhizopine catabolism (mocABRC) gene
A:Reference number: S51569; MUID:95147842
A:Accession: S51572
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-317 <ROS>
A:Cross-references: EMBL:X78503; NID:g468758; PIDN:CRA55269.1; PID:g468762
C:Genetics:
A:Gene: mocA
C:Superfamily: Streptomyces griseus strI protein

Query Match 4.1%; Score 7; DB 2; Length 317;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 ARGGAPR 152
|||||
DB 277 ARGGAPR 283

RESULT 16
A72065
GTP binding protein - Chlamydomophila pneumoniae (strain CWL029)
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000
C:Accession: A72065
R:Kallman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.; Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606
A:Accession: A72065
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-335 <ARN>
A:Cross-references: GB:AE001363; PIDN:AAD18684.1; PID:g4376831
A:Experimental source: strain CWL029
C:Genetics:
A:Gene: yhbZ
C:Superfamily: GTP-binding protein obg; translation elongation factor Tu homology
F:159-288/Domain: translation elongation factor Tu homology <ETU>
F:165-172/Region: nucleotide-binding motif A (P-loop)
F:285-288/Region: GTP-binding NKXD motif

Query Match 4.1%; Score 7; DB 2; Length 335;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 RDAETGE 45
|||||
DB 96 RDAETGE 102

RESULT 17
T26399
hypothetical protein Y105C5B.w - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
C:Accession: T26399
R:McMurray, A.
submitted to the EMBL Data Library, September 1999
A:Reference number: Z20208
A:Accession: T26399
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-338 <WIL>
A:Cross-references: EMBL:AL110479; NID:e1542153; PIDN:CAB54371.1; CESP:Y105C5B.w
A:Experimental source: clone Y105C5B
C:Genetics:
A:Gene: CESP:Y105C5B.w
A:Introns: 15/3; 51/2; 85/1; 170/2; 201/3; 268/3; 289/2
C:Superfamily: Caenorhabditis elegans hypothetical protein K06H6.2

Query Match 4.1%; Score 7; DB 2; Length 338;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 SLCLVL 16
|||||
DB 30 SLCLVL 36

RESULT 18
H72115
cationic amino acid transporter - Chlamydomophila pneumoniae (strains CWL029 and AR39)
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C:Accession: H72115; B81554
R:Kallman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.; Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606
A:Accession: H72115
A:Molecule type: DNA
A:Residues: 1-338 <ARN>
A:Cross-references: GB:AE001599; GB:AE001363; NID:g4376387; PIDN:AAD18280.1; PID:g437
A:Experimental source: strain CWL029
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39
A:Reference number: A81500; MUID:20150255
A:Accession: E81554
A:Molecule type: DNA
A:Residues: 1-338 <REA>
A:Cross-references: GB:AE002222; GB:AE002161; NID:g7189553; PIDN:AAF38460.1; PID:g718
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: ytfF; CP0645

Query Match 4.1%; Score 7; DB 2; Length 338;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 SLCLVL 16

Db 313 SLCLVL 319
|||||||

RESULT 19

G81602
GTP1/OBG family protein CP0208 [imported] - Chlamydomophila pneumoniae (strain AR39)
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 21-Jul-2000
C:Accession: G81602
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255
A:Accession: G81602
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-343 <REA>
A:Cross-references: GB:AE002181; GB:AE002161; NID:g7189129; PIDN:AAF38079.1; PID:g718913
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: CP0208
C:Superfamily: GTP-binding protein obg; translation elongation factor Tu homology

Query Match 4.1%; Score 7; DB 2; Length 343;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 RDAETGE 45
|||||||
Db 104 RDAETGE 110

RESULT 20

B55733
G protein-coupled receptor GPR2 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 26-Aug-1999
C:Accession: B55733
R:Marchese, A.; Docherty, J.M.; Nguyen, T.; Heiber, M.; Cheng, R.; Heng, H.H.Q.; Tsui, I
Genomics 23, 609-618, 1994
A:Title: Cloning of human genes encoding novel G protein-coupled receptors.
A:Reference number: A55733; MUID:95154831
A:Accession: B55733
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-354 <MAR>
A:Cross-references: GB:U13667
C:Genetics:
A:Gene: GDB:GPR2

A:Cross-references: GDB:371708; OMIM:600240
A:Map position: 17q21.1-17q21.3
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor

Query Match 4.1%; Score 7; DB 2; Length 354;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LALPALL 22
|||||||
Db 164 LALPALL 170

RESULT 21

JE0292
fungal stress protein - Rhizopus nigricans
C:Species: Rhizopus nigricans
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: JE0292

R:Cresnar, B.; Plaper, A.; Breskvar, K.; Hudnik-Plevnik, T.
Biochem. Biophys. Res. Commun. 250, 664-667, 1998
A:Title: cDNA sequence and deduced amino acid sequence of a fungal stress protein ind
A:Reference number: JE0292; MUID:99003515

A:Accession: JE0292
A:Molecule type: mRNA
A:Residues: 1-364 <CRE>
A:Cross-references: GB:Y10414
C:Superfamily: aldose 1-epimerase

Query Match 4.1%; Score 7; DB 1; Length 364;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 122 HAGFCLE 128
|||||||
Db 320 HAGFCLE 326

RESULT 22

E72549
probable acyl-CoA dehydrogenase APE1681 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: E72549
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Ta
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero
A:Reference number: A72450; MUID:99310339
A:Accession: E72549
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-393 <KAW>
A:Cross-references: DDBJ:AP000062; NID:g5105244; PIDN:BAAR0682.1; PID:g5105369
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1681
C:Superfamily: acyl-CoA dehydrogenase

Query Match 4.1%; Score 7; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GPGLSLL 12
|||||||
Db 71 GPGLSLL 77

RESULT 23

E71665
bicyclomycin resistance protein (bcr1) RP603 - Rickettsia prowazekii
C:Species: Rickettsia prowazekii
C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 20-Jun-2000
C:Accession: E71665
R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark
Nature 396, 133-140, 1998
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A:Reference number: A71630; MUID:99039499
A:Accession: E71665
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-407 <AND>
A:Cross-references: GB:AJ235272; GB:AJ235269; NID:g3861033; PIDN:CAAI5047.1; PID:g386
A:Experimental source: strain Madrid E
C:Genetics:
A:Gene: bcr1; RP603
C:Superfamily: bicyclomycin resistance protein

Query Match 4.1%; Score 7; DB 2; Length 407;

Best Local Similarity 100.0%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 SLCLVL 16
DB 378 SLCLVL 384
|||||

RESULT 24
T47281
hypothetical protein F26B15.30 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47281
R:Nyakatura, G.; Fartmann, B.; Dauner, D.; Sterr, W.; Holland, R.; Weichselgartner, M.;
Mayer, K.F.X.
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z24458
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-414 <NYA>
A:Cross-references: EMBL:AL138645
A:Experimental source: cultivar Columbia; BAC clone F26B15
C:Genetics:
A:Map position: 3
A:Introns: 21/3; 157/3; 208/3; 287/3; 319/3
A:Note: F26B15.30

Query Match 4.1%; Score 7; DB 2; Length 414;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 GEREEA 102
DB 56 GEREEA 62
|||||

RESULT 25
T35527
hypothetical protein SC6G10.18 - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 18-Aug-2000
C:Accession: T35527
R:Seeger, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, March 1999
A:Reference number: Z21581
A:Accession: T35527
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-432 <SEE>
A:Cross-references: EMBL:AL049497; PIDN:CAB39872.1; GSPDB:GN00070; SCOEDB:SC6G10.18
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SC6G10.18
C:Superfamily: yhaD protein

Query Match 4.1%; Score 7; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 NVLCGR 98
DB 204 NVLCGR 210
|||||

RESULT 26
T36609
probable secreted kinase - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T36609
R:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1999
A:Reference number: Z21610
A:Accession: T36609
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-433 <OLI>
A:Cross-references: EMBL:AL078610; PIDN:CAB44377.1; GSPDB:GN00070; SCOEDB:SCH35.50c
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SCH35.50c

Query Match 4.1%; Score 7; DB 2; Length 433;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 PALLPVP 25
DB 47 PALLPVP 53
|||||

RESULT 27
S29612
pectate lyase (EC 4.2.2.2) - trumpet lily
C:Species: Liliun longiflorum (trumpet lily)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 29-Sep-1999
C:Accession: S29612
R:Klm, S.; Finkel, D.J.; An, G.
submitted to the EMBL Data Library, October 1992
A:Description: Abundance patterns of lily pollen cDNAs: characterization of three pol
A:Reference number: S29611
A:Accession: S29612
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-434 <KIM>
A:Cross-references: EMBL:Z17328; NID:gl9450; PIDN:CAA78976.1; PID:gl9451
A:Experimental source: cv. Nellie White, mature flower
C:Superfamily: pectate lyase LAT59
C:Keywords: carbon-oxygen lyase

Query Match 4.1%; Score 7; DB 2; Length 434;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 PGTFVQR 60
DB 412 PGTFVQR 418
|||||

RESULT 28
H70333
conserved hypothetical protein aq_378 - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C:Accession: H70333
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: AY0300; MUID:98196666
A:Accession: H70333
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-453 <AQF>
A:Cross-references: GB:AE000687; NID:g2983050; PIDN:AAC06670.1; PID:g2983062; GB:AE00
A:Experimental source: strain VF5
C:Genetics:
A:Gene: aq_378

Query Match 4.1%; Score 7; DB 2; Length 453;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 FFAHAGF 125
|||||||
DB 406 FFAHAGF 412

RESULT 29

B83213

probable MFS transporter PA3467 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 15-Sep-2000
C:Accession: B83213
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Boman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: A82950
A:Accession: B83213
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-455 <STO>
A:Cross-references: GB:AE004767; GB:AE004091; NID:g9949604; PIDN:AAG06855.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA3467

Query Match

Best Local Similarity 4.1%; Score 7; DB 2; Length 455;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 LVLALPA 20
|||||||
DB 181 LVLALPA 187

RESULT 30

T22188

hypothetical protein F44F4.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T22188

R:Colles, L.
submitted to the EMBL Data Library, September 1994

A:Reference number: Z19528

A:Accession: T22188

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-472 <SWIL>

A:Cross-references: EMBL:Z37092; PIDN:CAA85457.1; GSPDB:GN000020; CESP:F44F4.6

A:Experimental source: clone F44F4

C:Genetics:

A:Gene: CESP:F44F4.6

A:Map position: 2

A:Introns: 57/2; 105/2; 185/3; 261/3; 319/1; 370/1; 412/3

Query Match

Best Local Similarity 4.1%; Score 7; DB 2; Length 472;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 PALLPVP 25
|||||||
DB 317 PALLPVP 323

RESULT 31

Q0ECRS

ygjE protein - Escherichia coli

C:Species: Escherichia coli
C:Date: 30-Jun-1988 #sequence_revision 31-Oct-1997 #text_change 22-Jun-1999
C:Accession: E65094; C29049
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; .A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617
A:Accession: E65094
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-487 <BLAT>
A:Cross-references: GB:A000388; GB:U00096; NID:gl789441; PIDN:AAC76099.1; PID:gl7894
A:Experimental source: strain K-12, substrain MG1655
R:Nesin, M.; Lupski, J.R.; Svec, P.; Godson, G.N.
Gene 51, 149-161, 1987
A:Title: Possible new genes as revealed by molecular analysis of a 5-kb Escherichia c
A:Reference number: A91573; MUID:87248073
A:Accession: C29049
A:Molecule type: DNA
A:Residues: 279-403, 'p', 405-411, 'RWRCRKSRCRCSA' <NES>
C:Genetics:
A:Gene: ygjE

A:Map position: 67 min
C:Superfamily: 2-oxoglutarate/malate translocator

C:Keywords: transmembrane protein
F:11-27/Domain: transmembrane #status predicted <TM1>
F:33-49/Domain: transmembrane #status predicted <TM2>
F:52-68/Domain: transmembrane #status predicted <TM3>
F:95-111/Domain: transmembrane #status predicted <TM4>
F:138-154/Domain: transmembrane #status predicted <TM5>
F:206-222/Domain: transmembrane #status predicted <TM6>
F:237-263/Domain: transmembrane #status predicted <TM7>
F:289-305/Domain: transmembrane #status predicted <TM8>
F:310-326/Domain: transmembrane #status predicted <TM9>
F:378-394/Domain: transmembrane #status predicted <TM10>
F:422-438/Domain: transmembrane #status predicted <TM11>
F:464-480/Domain: transmembrane #status predicted <TM12>

Query Match 4.1%; Score 7; DB 1; Length 487;

Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 ALLPVP 26
|||||||
DB 21 ALLPVP 27

RESULT 32

S08059

alpha-globulin type B precursor (tandem 1) - upland cotton (fragment)

N:Alternate names: seed storage protein

C:Species: Gossypium hirsutum (upland cotton)

C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 30-Sep-1993

C:Accession: S08059

R:Chian, C.A.; Borroto, K.; Kamalay, J.A.; Dure III, L.

Plant Mol. Biol. 9, 533-546, 1987

A:Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX.

A:Reference number: S06398

A:Accession: S08059

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-509 <CHL>

C:Superfamily: glycine

Query Match

Best Local Similarity 4.1%; Score 7; DB 2; Length 509;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 GEREEEA 102
|||||||

Db 78 GEREEA 84

RESULT 33

T45868

hypothetical protein F3A4.230 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C>Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000

C:Accession: T45868

R:Bargues, M.; Collado, M.C.; Navarro, P.; Terol, J.; Perez-Alonso, M.; Mewes, H.W.; May

submitted to the Protein Sequence Database, December 1999

A:Reference number: Z23007

A:Accession: T45868

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-509 <BAR>

A:Cross-references: EMBL:AL132978

A:Experimental source: cultivar Columbia; BAC clone F3A4

C:Genetics:

A:Map position: 3

A:Introns: 202/2; 369/1

A:Note: F3A4.230

Query Match

Best Local Similarity 4.1%; Score 7; DB 2; Length 509;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 EEEARAC 105

|||||

Db 157 EEEARAC 163

RESULT 34

T31115

histidine kinase homolog sdek - Myxococcus xanthus

C:Species: Myxococcus xanthus

C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999

C:Accession: T31115

R:Garza, A.G.; Pollack, J.S.; Harris, B.Z.; Lee, A.; Keseler, I.; Licking, E.F.; Singer,

submitted to the EMBL Data Library, October 1997

A:Description: A histidine kinase is required early in fruiting body development in myxo

A:Reference number: Z20990

A:Accession: T31115

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-513 <GAR>

A:Cross-references: EMBL:AF031084; NID:g2736295; PID:g2736297; PIDN:AAE94135.1

C:Genetics:

A:Gene: sdek

Query Match

Best Local Similarity 4.1%; Score 7; DB 2; Length 513;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 EREEAR 103

|||||

Db 269 EREEAR 275

RESULT 35

T47018

hypothetical protein [imported] - Yersinia pestis

C:Species: Yersinia pestis

C>Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000

C:Accession: T47018

R:Buchrieser, C.; Rusniok, C.; Couve, E.; Frangeul, L.; Billault, A.; Kunst, F.; Carniel

submitted to the EMBL Data Library, October 1998

A:Description: DNA sequence of the 102 kbases unstable region of Yersinia pestis.

A:Reference number: Z24348

A:Accession: T47018

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-546 <BUC>

A:Cross-references: EMBL:AL031866; PIDN:CAA21361.1

A:Experimental source: strain 6/69

Query Match

Best Local Similarity 4.1%; Score 7; DB 2; Length 546;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 RGVATP 34

|||||

Db 448 RGVATP 454

RESULT 36

S76764

hypothetical protein - Synechocystis sp. (strain PCC 6803)

C:Species: Synechocystis sp.

A:Variety: PCC 6803

C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999

C:Accession: S76764

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,

o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys

S.

A:Reference number: S74322; MUID:97061201

A:Accession: S76764

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-600 <KAN>

A:Cross-references: EMBL:D90916; GB:AB001339; NID:g1653715; PIDN:BAAL8676.1; PID:dl01

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Superfamily: aspartate kinase homology

F;3-417/Domain: aspartate kinase homology <DKI>

Query Match

Best Local Similarity 4.1%; Score 7; DB 2; Length 600;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 PAVRGVA 31

|||||

Db 433 PAVRGVA 439

RESULT 37

T03800

anthranilate synthase - Streptomyces violaceus

C:Species: Streptomyces violaceus

C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 11-May-2000

C:Accession: T03800

R:Lin, C.; Paraskar, A.S.; Vining, L.C.

Microbiology 144, 1971-1980, 1998

A:Title: Regulation of an anthranilate synthase gene in Streptomyces venezuelae by a

A:Reference number: Z15094; MUID:98361043

A:Accession: T03800

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-623 <LIN>

A:Cross-references: EMBL:AF012627; NID:g2318111; PID:g2318113

A:Experimental source: strain ISP5230

A:Note: the source is designated as Streptomyces venezuelae

C:Genetics:

A:Gene: trpE(G)

Query Match

Best Local Similarity 4.1%; Score 7; DB 2; Length 623;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 RDAETGE 45

Db 580 RDAETGE 586
|||||||

RESULT 38

S67175 probable membrane protein YOR273c - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein O5440

C:Species: Saccharomyces cerevisiae

C>Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 21-Jul-2000

C:Accession: S67175; S72044

R:Cheret, G.; Sor, F.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S67169

A:Accession: S67175

A:Molecule type: DNA

A:Residues: 1-659 <CHE>

A:Cross-references: EMBL:275181; NID:g1420611; PID:e252120; PID:g1420612; MIPS:YOR273c

A:Experimental source: strain S288C

R:Cheret, G.; Bernardi, A.; Sor, F.

Yeast 12, 1059-1064, 1996

A:Title: DNA sequence analysis of the VPH1-SNF2 region on chromosome XV of Saccharomyces

A:Reference number: S72039; MUID:97051594

A:Accession: S72044

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-659 <CHW>

A:Cross-references: EMBL:X89633; NID:g1279694; PIDN:CAA61779.1; PID:g1279700

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995

C:Genetics:

A:Map position: 15R

A:Note: YOR273c

C:Keywords: transmembrane protein

F:100-116/Domain: transmembrane #status predicted <TM1>

F:129-145/Domain: transmembrane #status predicted <TM2>

F:160-176/Domain: transmembrane #status predicted <TM3>

F:182-198/Domain: transmembrane #status predicted <TM4>

F:250-266/Domain: transmembrane #status predicted <TM5>

F:318-334/Domain: transmembrane #status predicted <TM6>

F:360-376/Domain: transmembrane #status predicted <TM7>

F:424-440/Domain: transmembrane #status predicted <TM8>

F:459-475/Domain: transmembrane #status predicted <TM9>

F:522-538/Domain: transmembrane #status predicted <TM10>

Query Match 4.1%; Score 7; DB 2; Length 659;

Best Local Similarity 100.0%; Pred. No. 63;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 RDAETGE 45

|||||||

Db 409 RDAETGE 415

RESULT 39

S52317

quinohemoprotein ethanol dehydrogenase (EC 1.1.99.-) type 1 precursor - Comamonas testos

C:Species: Comamonas testosteroni

C>Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 02-Jun-2000

C:Accession: S62366; S62373; S65908; S52317

R:Stoorvogel, J.; Kraayveld, D.E.; van Sluis, C.A.; Jongejan, J.A.; de Vries, S.; Duine,

Eur. J. Biochem. 235, 690-698, 1996

A:Title: Characterization of the gene encoding quinohemoprotein ethanol dehydrogenase d

A:Reference number: S62366; MUID:96184549

A:Accession: S62366

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-708 <STO1>

A:Cross-references: EMBL:X81880; NID:g663195; PIDN:CAA57464.1; PID:g663196

A:Accession: S62373

A:Status: preliminary

A:Molecule type: protein

A:Residues: 32-54 <STO2>

R;de Jong, G.A.H.; Geerlof, A.; Stoorvogel, J.; Jongejan, J.A.; de Vries, S.; Duine,
Eur. J. Biochem. 230, 899-905, 1995
A:Title: Quinohemoprotein ethanol dehydrogenase from Comamonas testosteroni. Purific
A:Reference number: S65908; MUID:95324580

A:Accession: S65908

A:Molecule type: protein

A:Residues: 32-50, 'X', 52-54; 477-483, 'X', 485-490 <DEJ>

A:Experimental source: ATCC 15667

C:Genetics:

A:Gene: qhdh

C:Superfamily: glucose dehydrogenase (pyrroloquinoline-quinone)

C:Keywords: oxidoreductase; quinoprotein

F:1-31/Domain: signal sequence #status predicted <SIG>

F:32-708/Product: quinohemoprotein ethanol dehydrogenase type 1 #status experimental

Query Match 4.1%; Score 7; DB 2; Length 708;

Best Local Similarity 100.0%; Pred. No. 67;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 DAETGER 46

|||||||

Db 227 DAETGER 233

RESULT 40

A83160

conserved hypothetical protein PA3893 [Imported] - Pseudomonas aeruginosa (strain PAO

C:Species: Pseudomonas aeruginosa

C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 15-Sep-2000

C:Accession: A83160

R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pa

A:Reference number: A82950

A:Accession: A83160

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-731 <STO>

A:Cross-references: GB:AE004806; GB:AE004091; NID:g9950067; PIDN:AAG07280.1; GSPDB:GN

A:Experimental source: strain PAO1

C:Genetics:

A:Gene: PA3893

Query Match 4.1%; Score 7; DB 2; Length 731;

Best Local Similarity 100.0%; Pred. No. 68;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 LCLVLVA 17

|||||||

Db 475 LCLVLVA 481

RESULT 41

E64807

outer membrane usher protein ybgQ precursor - Escherichia coli

C:Species: Escherichia coli

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: E64807

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;

.A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617

A:Accession: E64807

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-818 <BLAT>

A:Cross-references: GB:AE000175; GB:U00096; NID:g1786934; PIDN:AAC73812.1; PID:g17869

A:Experimental source: strain K-12, substrain MGL655

C:Genetics:
A:Gene: ybqQ
C:Function:
A:Description: biogenesis of fimbriae
C:Superfamily: outer membrane usher protein fimD
C:Keywords: membrane protein
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-818/Product: outer membrane usher protein ybqQ #status predicted <MAT>

Query Match 4.1%; Score 7; DB 1; Length 818;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 GVIAPGE 143
|||||
DB 49 GVIAPGE 55

RESULT 42
S18942
hypothetical protein (whiB 3' region) - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 22-Oct-1999
C:Accession: S18942; S20913
R:Davis, N.K.; Chater, K.F.
submitted to the EMBL Data Library, September 1991
A:Reference number: S18941
A:Accession: S18942
A:Molecule type: DNA
A:Residues: 1-914 <DAVI>
A:Cross-references: EMBL:X62287; NID:g49001; PIDN:CAA44176.1; PID:g49003
R:Davis, N.K.; Chater, K.F.
Mol. Gen. Genet. 232, 351-358, 1992
A:Title: The Streptomyces coelicolor whiB gene encodes a small transcription factor-like
A:Reference number: S20912; MUID:92269753
A:Accession: S20913
A:Molecule type: DNA
A:Residues: 1-11 <DAV2>
A:Cross-references: EMBL:X62287

Query Match 4.1%; Score 7; DB 2; Length 914;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 ALLPVA 26
|||||
DB 477 ALLPVA 483

RESULT 43
JE0363
mitogen-activated protein kinase kinase kinase (EC 2.7.-.-) - human
C:Species: Homo sapiens (man)
C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 11-May-2000
C:Accession: JE0363
R:Wang, X.; Diener, K.; Tan, T.; Yao, Z.
Biochem. Biophys. Res. Commun. 253, 33-37, 1998
A:Title: MAPKK6, a novel mitogen-activated protein kinase kinase kinase, that associates with mitogen-activated protein kinase
A:Reference number: JE0363; MUID:99092374
A:Accession: JE0363
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1288 <WAN>
A:Cross-references: GB:AF100318
C:Keywords: phosphotransferase

Query Match 4.1%; Score 7; DB 2; Length 1288;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 ETGERLV 48
|||||
DB 647 ETGERLV 653

RESULT 44
A55138
sodium channel mRNA2.3, voltage-gated - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Aug-1999
C:Accession: A55138
R:Feilipe, A.; Knattlie, T.J.; Doyle, K.L.; Tamkun, M.M.
J. Biol. Chem. 269, 30125-30131, 1994
A:Title: Primary structure and differential expression during development and pregnancy
A:Reference number: A55138; MUID:95074002
A:Accession: A55138
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-1681 <PEL>
A:Cross-references: GB:L36179; NID:g609544; PIDN:AAA66192.1; PID:g806397
C:Superfamily: sodium channel protein

Query Match 4.1%; Score 7; DB 2; Length 1681;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LALPALL 22
|||||
DB 1375 LALPALL 1381

RESULT 45
T30826
nascent polypeptide-associated complex alpha chain, muscle splice form gp220 - mouse
N:Alternate names: alpha-NAC protein
C:Species: Mus musculus (house mouse)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 24-Nov-1999
C:Accession: T30826
R:Yotov, W.V.; St-Arnaud, R.
Genes Dev. 10, 1763-1772, 1996
A:Title: Differential splicing-in of a proline-rich exon converts alphaNAC into a mus
A:Reference number: 220889; MUID:96312450
A:Accession: T30826
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2187 <YOT>
A:Cross-references: EMBL:U48363; NID:g1666688; PID:g1666689; PIDN:AAB18732.1
C:Genetics:
A:Gene: Naca
A:Map position: 10
A:Introns: 24/1; 1996/1; 2024/3; 2050/3; 2099/3; 2142/3; 2183/3
A:Note: differential splicing converts alphaNAC into a tissue-specific DNA-binding ac
C:Keywords: alternative splicing; DNA binding; transcription factor

Query Match 4.1%; Score 7; DB 2; Length 2187;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 AETPTYP 37
|||||
DB 974 AETPTYP 980

RESULT 46
LPHUB
apolipoprotein B-100 precursor - human
N:Contains: apolipoprotein B-26; apolipoprotein B-48; apolipoprotein B-74
C:Species: Homo sapiens (man)
C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 21-Jul-2000
C:Accession: A27850; A25679; A25263; A24320; A24684; A23817; A25774;
4452; I61909; I59510; I39474; I39469; I84624; I37179; PS0058

R:Ludwig, E.H.; Blackhart, B.D.; Pierotti, V.R.; Caiati, L.; Fortier, C.; Knott, T.; Soc DNA 6, 363-372, 1987
A:Title: DNA sequence of the human apolipoprotein B gene.
A:Reference number: A27850; MUID:88003974
A:Accession: A27850
A:Molecule type: mRNA
A:Residues: 1-617, 'A', 619-1929, 'F', 1931-3318, 'D', 3320-3426, 'T', 3428-3431, 'Q', 3433-3731, 'A'
A:Cross-references: GB:M14162
R:Cladaras, C.; Hadzopoulou-Cladaras, M.; Nolte, R.T.; Atkinson, D.; Zannis, V.I. EMBO J. 5, 3495-3507, 1986
A:Title: The complete sequence and structural analysis of human apolipoprotein B-100: re
A:Reference number: A91058; MUID:87161758
A:Accession: A25679
A:Molecule type: mRNA
A:Residues: 1-11, 15-2539, 'S', 2541-3823, 'R', 3825-4563 <CLA>
A:Note: 1109-Asp was also found
R:Knott, T.J.; Wallis, S.C.; Powell, L.M.; Pease, R.J.; Lusis, A.J.; Blackhart, B.; McCa Nucleic Acids Res. 14, 7501-7503, 1986
A:Title: Complete cDNA and derived protein sequence of human apolipoprotein B-100.
A:Reference number: A93639; MUID:87016385
A:Accession: A25263
A:Molecule type: mRNA
A:Residues: 1-272, 'N', 274-617, 'A', 619-1217, 'E', 1219-2091, 'V', 2093-2364, 'T', 2366-2679, 'Q'
A:Cross-references: GB:X04506; NID:g34330; PIDN:CAA28191.1; PID:g34331
R:Law, S.W.; Grant, S.M.; Higuchi, K.; Hospattankar, A.; Lackner, K.; Lee, N.; Brewer Jr Proc. Natl. Acad. Sci. U.S.A. 83, 8142-8146, 1986
A:Title: Human liver apolipoprotein B-100 cDNA: complete nucleic acid and derived amino
A:Reference number: A94134; MUID:87041416
A:Accession: A25267
A:Molecule type: mRNA
A:Residues: 1-617, 'A', 619-703, 'P', 705-792, 'R', 794-1270, 'S', 1272-1866, 'G', 1868-2036, 'N', 2
4189-4220, 'M', 4222-4563 <LAW>
A:Note: The codons given for residues 704, 793, 1271, 2037, 2933, 3286, 3782, 4188, and
R:Chen, S.H.; Yang, C.Y.; Chen, P.F.; Setzer, D.; Tanimura, M.; Li, W.H.; Gotto Jr., A.M J. Biol. Chem. 261, 12918-12921, 1986
A:Title: The complete cDNA and amino acid sequence of human apolipoprotein B-100.
A:Reference number: A92556; MUID:87008488
A:Accession: A25266
A:Molecule type: mRNA
A:Residues: 1-97, 'I', 99-328, 'V', 330-644, 'I', 646-918, 'P', 920-3318, 'D', 3320-3426, 'T', 3428-
9-4132, 'G', 4134-4180, 'E', 4182-4563 <CHE>
A:Cross-references: GB:X02610; NID:g178803; PIDN:AAA35549.1; PID:g178804
A:Note: A total of 2366 residues were confirmed by direct sequencing of tryptic peptides
R:Protter, A.A.; Hardman, D.A.; Sato, K.Y.; Schilling, J.W.; Yamanaka, M.; Hort, Y.J.; H Proc. Natl. Acad. Sci. U.S.A. 83, 5678-5682, 1986
A:Title: Analysis of cDNA clones encoding the entire B-26 region of human apolipoprotein
A:Reference number: A24320; MUID:86287319
A:Accession: A24320
A:Molecule type: mRNA
A:Residues: 1-97, 'I', 99-617, 'A', 619-941, 'Y', 941-1138, 'P', 1138-1138, 'PTGRLPNCFSNGLCYSLWLHSFQ
A:Cross-references: GB:M14081; NID:g178795; PIDN:AAA51752.1; PID:g553189
R:Law, S.W.; Lackner, K.J.; Hospattankar, A.V.; Anchors, J.M.; Sakaguchi, A.Y.; Naylor, Proc. Natl. Acad. Sci. U.S.A. 82, 8340-8344, 1985
A:Title: Human apolipoprotein B-100: cloning, analysis of liver mRNA, and assignment of
A:Reference number: A24684; MUID:86094221
A:Accession: A24684
A:Molecule type: mRNA
A:Residues: 485-617, 'A', 619-1044 <LA2>
A:Cross-references: GB:M12460; NID:g178791; PIDN:AAA51751.1; PID:g178792
R:Protter, A.A.; Hardman, D.A.; Schilling, J.W.; Miller, J.; Appleby, V.; Chen, G.C.; K Proc. Natl. Acad. Sci. U.S.A. 83, 1467-1471, 1986
A:Title: Isolation of a cDNA clone encoding the amino-terminal region of human apolipop
A:Reference number: A94088; MUID:86149325
A:Accession: A23817
A:Molecule type: mRNA
A:Residues: 1-291 <PRO>
A:Cross-references: GB:M12681; NID:g178787; PIDN:AAA51753.1; PID:g178798
R:Deeb, S.S.; Motulsky, A.G.; Albers, J.J. Proc. Natl. Acad. Sci. U.S.A. 82, 4983-4986, 1985
A:Title: A partial cDNA clone for human apolipoprotein B.
A:Reference number: A25774; MUID:85270450
A:Accession: A25774
A:Molecule type: mRNA

A:Residues: 709-791, 'SSWKAASHGCPHSAGD', 810-906 <DEE>
A:Cross-references: GB:X03175; NID:g178821; PIDN:AAA51759.1; PID:g178822
R:Carlsson, P.; Darnfors, C.; Olofsson, S.O.; Bjursell, G. Gene 49, 29-51, 1986
A:Title: Analysis of the human apolipoprotein B gene: complete structure of the B-74
A:Reference number: A91565; MUID:87191999
A:Accession: A26533
A:Molecule type: mRNA
A:Residues: 1282-2721, 2742-3290, 'L', 3292-3336, 'N', 3338-3948, 'F', 3950-3963, 'Y', 3965-41
A:Cross-references: GB:M15421; NID:g178817; PIDN:AAA51758.1; PID:g178818
R:Hardman, D.A.; Protter, A.A.; Chen, G.C.; Schilling, J.W.; Sato, K.Y.; Lau, K.; Yam Biochemistry 26, 5478-5486, 1987
A:Title: Structural comparison of human apolipoproteins B-48 and B-100.
A:Reference number: A29671; MUID:88050832
A:Accession: A29671
A:Molecule type: mRNA
A:Residues: 1671-2323, 'PYW', 2327-2352, 'H', 2354-2398 <HAR>
A:Cross-references: GB:M17367; NID:g178731; PIDN:AAA51741.1; PID:g178732
R:Shoulders, C.C.; Myant, N.B.; Sidoli, A.; Rodriguez, J.C.; Cortese, C.; Baralle, F. Atherosclerosis 58, 277-289, 1985
A:Title: Molecular cloning of human LDL apolipoprotein B cDNA. Evidence for more than
A:Reference number: A90084; MUID:86130855
A:Accession: A29287
A:Molecule type: mRNA
A:Residues: 3846-4298 <SHO>
R:Pfitzner, R.; Wagener, R.; Stoffel, W. Biol. Chem. Hoppe-Seyler 367, 1077-1083, 1986
A:Title: Isolation, expression and characterization of a human apolipoprotein B 100-S
A:Reference number: A25572; MUID:87076044
A:Accession: A25572
A:Molecule type: mRNA
A:Residues: 4219-4337, 'S', 4339-4563 <PFI>
A:Cross-references: GB:M36676
R:Wei, C.F.; Chen, S.H.; Yang, C.Y.; Marcel, Y.L.; Milne, R.W.; Li, W.H.; Sparrow, J. Proc. Natl. Acad. Sci. U.S.A. 82, 7265-7269, 1985
A:Reference number: A24738; MUID:86042646
A:Accession: A24738
A:Molecule type: mRNA
A:Residues: 'N', 3729-3731, 'I', 3733-3875, 'A', 3877-3948, 'F', 3950-3963, 'Y', 3965-3982, 'S'
A:Cross-references: GB:M12413; NID:g178735; PIDN:AAA51742.1; PID:g178736
R:Chen, S.H.; Habib, G.; Yang, C.Y.; Gu, Z.W.; Lee, B.R.; Weng, S.; Silberman, S.R.; Science 238, 363-366, 1987
A:Title: Apolipoprotein B-48 is the product of a messenger RNA with an organ-specific
A:Reference number: A40133; MUID:88018019
A:Accession: A40133
A:Molecule type: mRNA
A:Residues: 2165-2179 <CHI>
A:Cross-references: GB:M18036; NID:g178799; PIDN:AAA51754.1; PID:g178800
A:Note: this mRNA includes the stop codon of the organ-specific mRNA for apoB48
A:Accession: A40133
A:Molecule type: protein
A:Residues: 51-75, 101-110, 129-139, 158-174, 197-207, 276-287, 298-304, 306-314, 526-532, 538
36, 1486-1498, 1537-1556, 1563-1572, 1601-1610, 1647-1661, 1697-1724, 1770-1781, 1859-1897, 19
A:Note: these fragments were derived from apoB48
R:Hardman, D.A.; Protter, A.A.; Schilling, J.W.; Kane, J.P. Biochem. Biophys. Res. Commun. 149, 1214-1219, 1987
A:Title: Carboxyl terminal analysis of human B-48 protein confirms the novel mechanis
A:Reference number: A28002; MUID:88106542
A:Accession: A28002
A:Molecule type: mRNA
A:Residues: 2129-2179, 2181-2235 <HA2>
A:Cross-references: GB:M18471
A:Experimental source: Intestine
A:Note: this mRNA from intestine includes a stop codon created by RNA editing in plac
R:Menraban, M.; Schumaker, V.N.; Fareed, G.C.; West, R.; Johnson, D.F.; Kirchgessner Nucleic Acids Res. 13, 6937-6953, 1985
A:Title: Human apolipoprotein B: identification of cDNA clones and characterization o
A:Reference number: A24269; MUID:86041888
A:Accession: A24269
A:Molecule type: mRNA
A:Residues: 3056-3159 <MEH>
A:Cross-references: GB:X03045; NID:g28783; PIDN:CAA26850.1; PID:g929609
R:Hospattankar, A.V.; Higuchi, K.; Law, S.W.; Meglin, N.; Brewer Jr., H.B.

Biochem. Biophys. Res. Commun. 148, 279-285, 1987
A:Title: Identification of a novel in-frame translational stop codon in human intestine
A:Reference number: A29659; MUID:88049670
A:Accession: A29659
A:Molecule type: mRNA
A:Residues: 2169-2179 <HOS>
A:Note: the sequence shown represents the carboxyl end of apolipoprotein B-48
A:Note: two RNA species, 14.1kb and 7.5kb in length, were isolated from the human intestine
ch encodes the 250K apob-48, CAA encoding 2180-gln is substituted by the stop codon TAA.
R:Yang, C.; Kim, T.W.; Wang, S.; Lee, B.; Yang, M.; Gotto Jr., A.M.
Proc. Natl. Acad. Sci. U.S.A. 87, 5523-5527, 1990
A:Title: Isolation and characterization of sulfhydryl and disulfide peptides of human ap
A:Reference number: A35783; MUID:90319144
A:Contents: disulfide bonds
A:Accession: A35783
A:Molecule type: protein
A:Residues: 28-41;76-97, 'R',99-100;175-193;206-215;239-249;259-266;357-399;455-490;512-5
A:Note: cysteines at positions 1112, 1422, 1505, 1662, 3761, 3917, and 4217 have free su
R:LeBeauf, R.C.; Miller, C.; Shively, J.E.; Schumaker, V.N.; Balla, M.A.; Lusis, A.J.
FEBS Lett. 170, 105-108, 1984
A:Title: Human apolipoprotein B: partial amino acid sequence.
A:Reference number: A22006; MUID:84208786
A:Accession: A22006
A:Molecule type: protein
A:Residues: 873-892, 'K',894-896 <LEI>
A:Accession: B22006
A:Molecule type: protein
A:Residues: 3113, 'L',3115-3130, 'P',3132-3133, 'P',3135-3136, 'R' <LE2>
R:Blackhart, B.D.; Ludwig, E.M.; Pierotti, V.R.; Caiati, L.; Onasch, M.A.; Wallis, S.C.;
J. Biol. Chem. 261, 15364-15367, 1986
A:Title: Structure of the human apolipoprotein B gene.
A:Reference number: A92564; MUID:87057153
A:Contents: annotation: gene structure
R:Wagner, R.; Pritzner, R.; Stoffel, W.
Biol. Chem. Hoppe-Seyler 368, 419-425, 1987
A:Title: Studies on the organization of the human apolipoprotein B 100 gene.
A:Reference number: A90715; MUID:87271140
A:Contents: annotation: gene structure
R:Weisgraber, K.H.; Rall Jr., S.C.
J. Biol. Chem. 262, 11097-11103, 1987
A:Title: Human apolipoprotein B-100 heparin-binding sites.
A:Reference number: A92605; MUID:87280197
A:Contents: annotation: heparin binding and disulfide bond
R:Dashti, N.; Lee, D.M.; Mok, T.
Biochem. Biophys. Res. Commun. 137, 493-499, 1986
A:Title: Apolipoprotein B is a calcium binding protein.
A:Reference number: A90125; MUID:86242245
A:Contents: annotation: calcium binding
R:Carlsson, P.; Olofsson, S.O.; Bondjers, G.; Darnfors, C.; Wiklund, O.; Bjursell, G.
Nucleic Acids Res. 13, 8813-8826, 1985
A:Title: Molecular cloning of human apolipoprotein B cDNA.
A:Reference number: I37178; MUID:86093680
A:Accession: I37180

Query Match 4.1%; Score 7; DB 1; Length 4563;
Best Local Similarity 100.0%; Pred. No. 2.7e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LALPALL 22
|
Db 12 LALPALL 18

RESULT 47
PC4437
hydrogenase (EC 1.18.99.1) (Nlfe) - Desulfovibrio vulgaris (fragment)
C:Species: Desulfovibrio vulgaris
C:Date: 04-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 07-May-1999
C:Accession: PC4437
R:Romao, C.V.; Pereira, I.A.C.; Xavier, A.V.; LeGall, J.; Teixeira, M.
Biochem. Biophys. Res. Commun. 240, 75-79, 1997
A:Title: Characterization of the [Nlfe] hydrogenase from the sulfate reducer Desulfovibrio

A:Reference number: PC4437; MUID:98042519
A:Accession: PC4437
A:Molecule type: protein
A:Residues: 1-25 <ROW>
C:Keywords: oxidoreductase

Query Match 3.5%; Score 6; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 AETGER 46
|
Db 15 AETGER 20

RESULT 48
T14905
hypothetical protein - parsley
C:Species: Petroselinum crispum (parsley)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T14905
R:Feldbrugge, M.; Sprenger, M.; Dinkelbach, M.; Yazaki, K.; Harter, K.; Weisshaar, B.
Plant Cell 6, 1607-1621, 1994
A:Title: Functional analysis of a light-responsive plant bZIP transcriptional regulat
A:Reference number: Z18259; MUID:95128172
A:Accession: T14905
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-28 <FE>
A:Cross-references: EMBL:S75395; NID:g913201; PID:e194726

Query Match 3.5%; Score 6; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 LVLALP 19
|
Db 5 LVLALP 10

RESULT 49
S78708
protein YBRI26w-a - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C:Date: 15-Jan-1999 #sequence_revision 15-Jan-1999 #text_change 15-Jan-1999
C:Accession: S78708
R:Becam, A.M.; Herbert, C.J.; Nasr, F.; Slonimski, P.P.; Zagulski, M.
submitted to the Protein Sequence Database, August 1994
A:Reference number: S45995
A:Accession: S78708
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-47 <BEC>
A:Cross-references: EMBL:Z35995; MIPS:YBRI26w-a
C:Genetics:
A:Map position: 2R

Query Match 3.5%; Score 6; DB 2; Length 47;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 90 YCNVLC 95
|
Db 4 YCNVLC 9

RESULT 50
S33413
NIMA protein - human
C:Species: Homo sapiens (man)

C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 07-May-1999
C;Accession: S33413
R;Tang, T.K.; Tang, C.C.; Chen, Y.L.; Wu, C.W.
J. Cell Sci. 104, 249-260, 1993
A;Title: Nuclear proteins of the bovine esophageal epithelium. II. The NuMA gene gives rise to two distinct protein products.
A;Reference number: S33376; MUID:93280231
A;Accession: S33413
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-58 <TAN>
A;Cross-references: EMBL:Z14228

Query Match 3.5%; Score 6; DB 2; Length 58;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 165 GPSLAP 170
|||||
Db 29 GPSLAP 34

Search completed: January 30, 2001, 17:08:02
Job time: 766 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 17:10:43 ; Search time 17.99 seconds
(without alignments)
305.169 Million cell updates/sec

Title: US-09-518-931-4
Perfect score: 170
Sequence: 1 MRALEGPGLSLCLVLALPA.....PRSGRRGCGVAGPGLAP 170

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 88757 seqs, 32294092 residues

Word size : 6

Total number of hits satisfying chosen parameters: 239

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	4.7	530	1 YDEV_ECOLI	P77432 escherichia
2	8	4.7	1172	1 TSP2_HUMAN	P35442 homo sapien
3	7	4.1	113	1 IMMO_ECOLI	P33637 escherichia
4	7	4.1	113	1 IMML_ECOLI	P02985 escherichia
5	7	4.1	113	1 IMML_SHISO	P22558 shigella so
6	7	4.1	317	1 MOCA_RHIME	P49307 rhizobium m
7	7	4.1	337	1 NOD2_BRASN	P50328 bradyrhizob
8	7	4.1	362	1 CKRA_HUMAN	P46092 homo sapien
9	7	4.1	434	1 PEL_LILLO	P40973 lillium long
10	7	4.1	482	1 RRB_MOUSE	P22605 mus musculus
11	7	4.1	487	1 TTDT_ECOLI	P39414 escherichia
12	7	4.1	535	1 MVIN_RHIME	P56882 rhizobium m
13	7	4.1	818	1 YBGQ_ECOLI	P75750 escherichia
14	7	4.1	903	1 ECLC_BOVIN	P54281 bos taurus
15	7	4.1	1011	1 M3K6_HUMAN	O95382 homo sapien
16	7	4.1	4563	1 APB_HUMAN	P04114 homo sapien
17	6	3.5	61	1 XVLB_ACTMI	P12867 actinoplan
18	6	3.5	88	1 VG30_BPMD2	O64224 mycobacteri
19	6	3.5	92	1 HRK_MOUSE	P70678 mus musculus
20	6	3.5	92	1 LE10_HELAN	P46514 heliathanth
21	6	3.5	98	1 FIXX_RHILT	P08710 rhizobium 1
22	6	3.5	102	1 YXCD_BACSU	P46334 bacillus su
23	6	3.5	110	1 NOLE_RHILP	P23716 rhizobium 1
24	6	3.5	111	1 SC4_SCHCO	P16934 schizophyll
25	6	3.5	125	1 NEUI_HUMAN	P01178 homo sapien
26	6	3.5	125	1 NEUI_PIG	P01177 sus scrofa
27	6	3.5	134	1 NB8M_HUMAN	P17568 homo sapien
28	6	3.5	137	1 MP12_LYMS	P25289 lymnaea sta
29	6	3.5	138	1 MP15_LYMS	P31241 lymnaea sta
30	6	3.5	143	1 PER_DROPI	Q25206 drosophila
31	6	3.5	150	1 CNRD_BOVIN	O95142 bos taurus
32	6	3.5	150	1 CNRD_CANFA	O9xt54 canis famil
33	6	3.5	150	1 CNRD_HUMAN	O43924 homo sapien

34	6	3.5	150	1 CNRD_MOUSE	O55057 mus musculus
35	6	3.5	159	1 YK91_CAPEL	Q18268 caenorhabdi
36	6	3.5	161	1 UL20_PRVN3	Q00702 pseudorabie
37	6	3.5	166	1 VPB_BPP2	P07636 bacterioph
38	6	3.5	168	1 YH11_AZOBR	P26723 azospirillu
39	6	3.5	171	1 YC50_AQUAE	O67293 aquifex aeo
40	6	3.5	172	1 YHL4_HCNVA	P09698 human cyt
41	6	3.5	189	1 YHNN_CAPEL	O94402 caenorhabdi
42	6	3.5	191	1 CLD7_RAT	O92111 rattus norv
43	6	3.5	191	1 YGL2_STRCO	P40182 streptomyce
44	6	3.5	198	1 YC69_PYRHO	O58984 pyrococcus
45	6	3.5	203	1 IDI1_MYCTU	Q63086 rattus norv
46	6	3.5	203	1 IDI1_MYCTU	P72002 mycobacteri
47	6	3.5	210	1 FTSQ_AGRTU	O30990 agrobacteri
48	6	3.5	210	1 LAG1_HUMAN	O75638 homo sapien
49	6	3.5	228	1 COX2_LUNTE	Q37545 lumbricus t
50	6	3.5	231	1 GPHC_ALCEU	P40852 alcaligenes
51	6	3.5	231	1 GPHC_ALCEU	P40853 alcaligenes
52	6	3.5	233	1 HEX8_ADEL2	P36713 human adeno
53	6	3.5	238	1 RLI1_RICPR	O9ze23 rickettsia
54	6	3.5	238	1 YPE2_RHURU	P23139 rhodospiril
55	6	3.5	241	1 DCOP_PYRHO	O58482 pyrococcus
56	6	3.5	244	1 ARC3_CBCP	Q00901 clostridium
57	6	3.5	247	1 GTS2_DROME	P41043 drosophila
58	6	3.5	251	1 ARC3_CBCP	P15879 clostridium
59	6	3.5	251	1 COBM_MYCTU	O10672 mycobacteri
60	6	3.5	253	1 Y883_MYCTU	O10545 mycobacteri
61	6	3.5	257	1 TDXM_BOVIN	P35705 bos taurus
62	6	3.5	257	1 VATD_CAEL	P34462 caenorhabdi
63	6	3.5	260	1 NRPN_RAT	Q88780 rattus norv
64	6	3.5	268	1 CEBD_MOUSE	Q00322 mus musculus
65	6	3.5	268	1 CEBD_RAT	O03484 rattus norv
66	6	3.5	276	1 MURI_LACBR	P48797 lactobacill
67	6	3.5	283	1 TPMM_LOCM1	P31816 locusta mig
68	6	3.5	287	1 FCAR_HUMAN	P24071 homo sapien
69	6	3.5	289	1 NOLT_RHIFR	P33209 rhizobium f
70	6	3.5	289	1 NOLT_RHISN	P55714 rhizobium s
71	6	3.5	292	1 Y125_MYCTU	O50608 mycobacteri
72	6	3.5	293	1 TYRC_ZYMO	Q04983 zymomonas m
73	6	3.5	293	1 VBL1_BGNV	P06001 bean golden
74	6	3.5	293	1 VBL1_PYMOV	P27267 potato yell
75	6	3.5	293	1 VBL1_SLCV	P21936 squash leaf
76	6	3.5	293	1 VBL1_TGNV	O35566 tomato gold
77	6	3.5	293	1 VBL1_TMOV	O36660 tomato mott
78	6	3.5	298	1 YC34_HELPY	O25832 helicobacte
79	6	3.5	300	1 YP79_MYCTU	O50642 mycobacteri
80	6	3.5	301	1 PHLP_RAT	Q63737 rattus norv
81	6	3.5	306	1 OPPB_ECOLI	P31132 escherichia
82	6	3.5	306	1 OPPB_HAEIN	P45054 haemophilus
83	6	3.5	306	1 OPPB_SALT	P08005 salmonella
84	6	3.5	312	1 CAH4_HUMAN	P22748 homo sapien
85	6	3.5	317	1 FCN2_HUMAN	O14257 homo sapien
86	6	3.5	319	1 PTHB_ECOLI	P56580 escherichia
87	6	3.5	319	1 PTHB_ECOLI	P23851 escherichia
88	6	3.5	320	1 MDCE_RHIME	P56949 rhizobium m
89	6	3.5	322	1 HYPE_ECOLI	P24193 escherichia
90	6	3.5	323	1 HFLC_BORBU	O51222 borrelia bu
91	6	3.5	325	1 DBP_MOUSE	Q50925 mus musculus
92	6	3.5	325	1 DBP_RAT	P16443 rattus norv
93	6	3.5	327	1 RL5_ANOGA	O44248 anopheles g
94	6	3.5	329	1 BCSP_BRUAB	P12920 brucella ab
95	6	3.5	333	1 PTHB_ERWAM	O32522 erwinia amy
96	6	3.5	336	1 PTHB_GLOBE	O32333 clostridium
97	6	3.5	344	1 ARC2_THETH	O50146 thermus aqu
98	6	3.5	352	1 YK50_YEAST	P36151 saccharomyc
99	6	3.5	356	1 SPAS_SALT	P40702 salmonella
100	6	3.5	360	1 Y871_RICPR	O9zcv95 rickettsia
101	6	3.5	361	1 PAX1_HUMAN	P15863 homo sapien
102	6	3.5	365	1 SUCC_METH	O27115 methanobact
103	6	3.5	368	1 HIP_RAT	P50503 rattus norv
104	6	3.5	369	1 HIP_HUMAN	P50502 homo sapien
105	6	3.5	370	1 GPRA_RAT	O64121 rattus norv
106	6	3.5	376	1 KAPR_CAEL	P30625 caenorhabdi

107	6	3.5	376	1	Y501_MYCTU	Q11166 mycobacteri	180	6	3.5	671	1	YPFI_ECOLI	P76562 escherichia
108	6	3.5	381	1	ME24_ASPFL	P46073 aspergillus	181	6	3.5	690	1	HRPO_BURSO	P35656 burkholderi
109	6	3.5	382	1	ACBC_ASPFL	Q92ae9 actinoplan	182	6	3.5	695	1	MDLI_YEAST	P33310 saccharomyc
110	6	3.5	384	1	DUS9_HUMAN	Q99956 homo sapien	183	6	3.5	730	1	COG9_MOUSE	P41245 mus musculus
111	6	3.5	388	1	CHS_VIGN	P31089 vigna ungui	184	6	3.5	731	1	DCOR_ECOLI	P21169 escherichia
112	6	3.5	388	1	DACD_ECOLI	P33013 escherichia	185	6	3.5	749	1	DNK3_SYN7	P50022 synecococc
113	6	3.5	392	1	GAG_BLVAV	P25058 bovine leuk	186	6	3.5	755	1	COMP_RAT	P35444 rattus norv
114	6	3.5	395	1	HMPA_ERWCH	Q47266 erwinia chr	187	6	3.5	764	1	CPT1_MOUSE	P97742 mus musculu
115	6	3.5	396	1	HMPA_ECOLI	P24232 escherichia	188	6	3.5	773	1	CPT1_HUMAN	P50416 homo sapien
116	6	3.5	396	1	HMPA_SALTY	P26353 salmonella	189	6	3.5	773	1	CPT1_RAT	P32198 rattus norv
117	6	3.5	398	1	GCH2_BACSU	P17620 bacillus su	190	6	3.5	809	1	BCSB_ACEXY	P37717 acetobacter
118	6	3.5	400	1	PCAF_PSEPU	Q51956 pseudomonas	191	6	3.5	845	1	PPSA_AERPE	O9vec5 aeropyrum p
119	6	3.5	406	1	IPSP_HUMAN	P05154 homo sapien	192	6	3.5	871	1	UL47_HSVEB	P28929 equine herp
120	6	3.5	406	1	IPSP_HUMAN	P32057 escherichia	193	6	3.5	886	1	ITH3_MOUSE	O63174 rattus norv
121	6	3.5	410	1	SEX3_MYCBO	O07129 mycobacteri	194	6	3.5	887	1	ITH3_RAT	P36617 schizosacch
122	6	3.5	410	1	SEX3_MYCTU	Q11155 mycobacteri	195	6	3.5	892	1	RA16_SCHPO	O00653 homo sapien
123	6	3.5	410	1	YG31_MYCLE	Q50178 mycobacteri	196	6	3.5	898	1	KBF2_HUMAN	P02730 homo sapien
124	6	3.5	419	1	FAAA_HUMAN	P16930 homo sapien	197	6	3.5	911	1	B3AT_HUMAN	P53814 homo sapien
125	6	3.5	426	1	DHMH_PARVE	P23006 paracoccus	198	6	3.5	915	1	SMOO_HUMAN	O43099 homo sapien
126	6	3.5	441	1	VG34_HSV1	Q00155 ictaluriid h	199	6	3.5	919	1	EXL3_HUMAN	O08387 saccharomyc
127	6	3.5	445	1	YGCS_ECOLI	O46909 escherichia	200	6	3.5	944	1	RHG4_YEAST	P98171 homo sapien
128	6	3.5	447	1	YML3_ARATH	O23693 arabidopsis	201	6	3.5	956	1	TSP3_HUMAN	P49746 homo sapien
129	6	3.5	448	1	PRTE_ERWCH	P23597 erwinia chr	202	6	3.5	956	1	TSP3_MOUSE	O05895 mus musculu
130	6	3.5	452	1	AMT_CORGL	P54146 corynebacte	203	6	3.5	975	1	KINH_DROME	P17210 drosophila
131	6	3.5	460	1	SUOX_CHICK	P07850 gallus gall	204	6	3.5	976	1	FIBP_ADEB3	Q03553 bovine aden
132	6	3.5	461	1	SSL1_YEAST	O04673 saccharomyc	205	6	3.5	976	1	YKIC2_CAEEL	P41993 caenorhabdi
133	6	3.5	462	1	ALGC_PSEAE	P26276 pseudomonas	206	6	3.5	1018	1	DSGL_HUMAN	O02413 homo sapien
134	6	3.5	467	1	EUTA_ECOLI	P76551 escherichia	207	6	3.5	1049	1	IF2P_SCHPO	O10251 schizosacch
135	6	3.5	467	1	EUTA_SALTY	Q92fiv2 salmonella	208	6	3.5	1079	1	TSP2_BOVIN	O95116 bos taurus
136	6	3.5	471	1	RBL2_HYDMR	Q59460 hydrogenovi	209	6	3.5	1170	1	TSP1_CHICK	P35440 gallus gall
137	6	3.5	472	1	ARAE_ECOLI	P09830 escherichia	210	6	3.5	1178	1	DNBI_HSV2	P89452 herpes simp
138	6	3.5	472	1	ARAE_KLEOX	P45598 klebsiella	211	6	3.5	1196	1	POLN_WEEV	P36384 herpes simp
139	6	3.5	473	1	RBL1_RHOCA	O32740 rhodobacter	212	6	3.5	1236	1	POLN_WEEV	P13897 western equ
140	6	3.5	473	1	RBL1_THIFE	P28895 thiobacillu	213	6	3.5	1239	1	POLN_WEEV	P08768 eastern equ
141	6	3.5	473	1	RBL2_THIFE	Q07087 thiobacillu	214	6	3.5	1240	1	POLN_WEEV	P27284 eastern equ
142	6	3.5	474	1	TED_HUMAN	O75949 homo sapien	215	6	3.5	1275	1	COBN_PSEDE	P29939 pseudomonas
143	6	3.5	474	1	RBL_PSEHY	Q51856 pseudomonas	216	6	3.5	1300	1	DYNA_NEUCR	Q01397 neurospora
144	6	3.5	483	1	PREG_NEUCR	O06712 neurospora	217	6	3.5	1300	1	IRR_CAVPO	P14617 cavia porce
145	6	3.5	491	1	GABD_RHISN	P55653 rhizobium s	218	6	3.5	1443	1	E75C_DROME	P13055 drosophila
146	6	3.5	497	1	DTPT_LACHE	O07380 lactobacill	219	6	3.5	1607	1	MYPR_LYMT	Q25410 lymnaea sta
147	6	3.5	498	1	IL14_HUMAN	P40222 homo sapien	220	6	3.5	1698	1	Y076_HUMAN	O14999 homo sapien
148	6	3.5	502	1	CP83_ARATH	P48421 arabidopsis	221	6	3.5	1857	1	MYST_HUMAN	P35749 homo sapien
149	6	3.5	502	1	K2C8_XENLA	P08776 xenopus lae	222	6	3.5	1900	1	MYST_HUMAN	P37297 saccharomyc
150	6	3.5	503	1	YKR5_YEAST	P34240 saccharomyc	223	6	3.5	1905	1	STT4_YEAST	P54683 dictyostell
151	6	3.5	516	1	LEGB_GOSHI	P09800 gossypium h	224	6	3.5	1972	1	MYST_RABIT	P35748 oryctolagus
152	6	3.5	520	1	RP54_RHIE	P49989 rhizobium e	225	6	3.5	1978	1	MYSG_CHICK	P10587 gallus gall
153	6	3.5	521	1	Y725_CAEEL	Q10934 caenorhabdi	226	6	3.5	2182	1	POLG_CXBLJ	P08291 c genome po
154	6	3.5	524	1	CPVL_ICTPU	Q92111 ictalurus p	227	6	3.5	2191	1	POLG_EC06C	O66474 e genome po
155	6	3.5	531	1	ACHC_ACHFU	P35903 achatina fu	228	6	3.5	2201	1	ABCI_HUMAN	O95477 homo sapien
156	6	3.5	532	1	ICAI_HUMAN	P03362 homo sapien	229	6	3.5	2201	1	ABCI_MOUSE	P41233 mus musculu
157	6	3.5	533	1	VMCB_LAMBD	P03710 bacterioph	230	6	3.5	2224	1	FA5_HUMAN	P12259 homo sapien
158	6	3.5	545	1	FGR_FSVGR	P00544 feline sarc	231	6	3.5	2286	1	DPOE_HUMAN	Q07864 homo sapien
159	6	3.5	547	1	CP78_MAIZE	P48420 zea mays (m	232	6	3.5	2301	1	POLG_TMEVD	P13899 t genome po
160	6	3.5	554	1	HYES_HUMAN	P34913 homo sapien	233	6	3.5	2303	1	POLG_TMEVB	O08544 t genome po
161	6	3.5	556	1	PMGI_RICCO	P35493 ricinus com	234	6	3.5	2647	1	ABP2_HUMAN	P12333 homo sapien
162	6	3.5	558	1	FLAJ_METVO	O59649 methanococc	235	6	3.5	2805	1	MAPA_HUMAN	P78559 homo sapien
163	6	3.5	559	1	PMGI_MESCR	Q42908 mesembryant	236	6	3.5	3033	1	POLG_HCVJ8	P26661 h genome po
164	6	3.5	581	1	IRR_RAT	Q64716 rattus norv	237	6	3.5	3707	1	PGBM_MOUSE	Q05793 mus musculu
165	6	3.5	587	1	LACP_KLULA	P07921 kluyveromyc	238	6	3.5	4393	1	PGBM_HUMAN	P98160 homo sapien
166	6	3.5	588	1	VCLA_GOSHI	P09801 gossypium h	239	6	3.5				
167	6	3.5	591	1	P091_YEAST	P54113 s bifunctio							
168	6	3.5	592	1	CEA_CITFR	P04480 citrobacter							
169	6	3.5	592	1	P092_YEAST	P38009 s bifunctio							
170	6	3.5	594	1	NIFA_RHISN	Q53206 rhizobium s							
171	6	3.5	594	1	NPAL_MOUSE	P97459 mus musculu							
172	6	3.5	595	1	RUBB_PEA	P08927 pisum sativ							
173	6	3.5	604	1	RCOL_NEUCR	P78706 neurospora							
174	6	3.5	617	1	VG_F_RAT	P20156 rattus norv							
175	6	3.5	622	1	NTSE_DROME	P51905 drosophila							
176	6	3.5	627	1	Y017_RICPR	Q92ec6 rickettsia							
177	6	3.5	629	1	YS50_MYCTU	O05809 mycobacteri							
178	6	3.5	631	1	DHML_PARDE	P12293 paracoccus							
179	6	3.5	664	1	CHM2_RABIT	Q28718 oryctolagus							

ALIGNMENTS

RESULT 1
YDEV_ECOLI STANDARD; PRT: 530 AA.
ID YDEV_ECOLI
AC P77432; Q99894;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL SUGAR KINASE IN HIPB-UXAB INTERGENIC REGION.

GN YDEV.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12";
 RL Science 277:1453-1474(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RA Alba H., Baba T., Fujita K., Hayashi K., Honjo A., Horiuchi T.,
 RA Ikemoto K., Inada T., Isono K., Itoh T., Kanai K., Kasai H.,
 RA Kashimoto K., Kim S., Kimura S., Kitagawa M., Kitakawa M., Makino K.,
 RA Masuda S., Miki T., Mizobuchi K., Mori H., Motomura K., Nakamura Y.,
 RA Nishimoto H., Nishio Y., Oshima T., Saito N., Sampaio G., Seki Y.,
 RA Tagami H., Takemoto K., Wada C., Yamamoto Y., Yano M.;
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 182-495 FROM N.A.
 RX MEDLINE=96243037; PubMed=8649811;
 RA Das R., Reddy E.P., Chatterjee D., Andrews D.W.;
 RT "Identification of a novel Bcl-2 related gene, BRAG-1, in human
 RT glioma";
 RL Oncogene 12:947-951(1996).
 CC -1- SIMILARITY: BELONGS TO THE FUCOKINASE / GLUCOKINASE /
 CC GLYCOKINASE / XYLOKINASE FAMILY.
 CC -1- CAUTION: WAS THOUGHT BY REF.3 TO BE A HUMAN SEQUENCE AND WAS
 CC WITH A ROLE IN APOPTOSIS. THE DNA SEQUENCE OF THE REGION THEY
 CC SEQUENCED IS MORE THAN 99% IDENTICAL TO THAT OF THIS E.COLI
 CC GENE. FURTHERMORE THEY CLAIM 'EXTENSIVE SIMILARITY TO THE
 CC BCL-2 FAMILY OF GENES.' SUCH A SIMILARITY IS NOT SIGNIFICANT
 CC AND THIS PROTEIN IS MUCH MORE LIKELY TO BE A SUGAR KINASE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE000248; AAC74584.1; -;
 DR EMBL; D90793; BAA15191.1; -;
 DR EMBL; D90794; BAA15198.1; -;
 DR EMBL; S82185; AAC17184.1; -;
 DR ECOLINE; EGI3804; YDEV.
 DR INTERPRO; IPR000577; -;
 DR PFAM; PF00370; FGGY; 2.
 DR PROSITE; PS00933; FGGY_KINASES_1; FALSE_NEG.
 DR PROSITE; PS00445; FGGY_KINASES_2; FALSE_NEG.
 KW Hypothetical protein; Transferase; Kinase.
 FT CONFLICT 490 495 PDPEKH -> TRPGKA (IN REF. 2).
 SQ SEQUENCE 530 AA; 57544 MW; C8C3B1E7C8982063 CRC64;

Query Match 4.7%; Score 8; DB 1; Length 530;
 Best Local Similarity 100.0%; Pred. No. 2.7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 41 AETGERLV 48
 |||||
 Db 475 AETGERLV 482

RESULT 2

TSP2_HUMAN
 ID TSP2_HUMAN STANDARD; PRT; 1172 AA.
 AC P35442;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE THROMBOSPONDIN 2 PRECURSOR.
 GN THBS2 OR TSP2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94010892; PubMed=8406456;
 RA Labell T.L., Byers P.H.;
 RT "Sequence and characterization of the complete human thrombospondin 2
 RT cDNA: potential regulatory role for the 3' untranslated region.";
 RL Genomics 17:225-229(1993).
 RN [2]
 RP SEQUENCE OF 560-1172 FROM N.A.
 RX TISSUE-FIBROBLAST;
 RA MEDLINE=92217961; PubMed=1559694;
 RA Labell T.L., McGookey Milewicz D.J., Distche C.M., Byers P.H.;
 RT "Thrombospondin II: partial cDNA sequence, chromosome location, and
 RT expression of a second member of the thrombospondin gene family in
 RT humans";
 RL Genomics 12:421-429(1992).
 CC -1- FUNCTION: ADHESIVE GLYCOPROTEIN THAT MEDIATES CELL-TO-CELL AND
 CC CELL-TO-MATRIX INTERACTIONS. CAN BIND TO FIBRINOGEN, FIBRONECTIN,
 CC LAMININ AND TYPE V COLLAGEN.
 CC -1- SUBUNIT: HOMOTRIMER, CROSS-LINKED BY DISULFIDE BONDS.
 CC -1- SIMILARITY: BELONGS TO THE THROMBOSPONDIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 VMFC DOMAIN.
 CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 3 TYPE-1 TSP REPEATS AND 7 TYPE-3 TSP REPEATS
 CC (WHICH BIND CALCIUM).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L12350; AAA03703.1; -;
 DR EMBL; M81339; -; NOT_ANNOTATED_CDS.
 DR PIR; A42173; A42173.
 DR HSSP; P00740; 11XA.
 DR MIN; 188061; -;
 DR INTERPRO; IPR000561; -;
 DR INTERPRO; IPR000884; -;
 DR INTERPRO; IPR001007; -;
 DR PFAM; PF00008; EGF; 2.
 DR PFAM; PF00090; tsp_1; 3.
 DR PFAM; PF00093; vwc; 1.
 DR PROSITE; PS00022; EGF_1; FALSE_NEG.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS00092; TSP1; 3.
 DR PROSITE; PS01208; VMFC; 1.
 DR Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
 DR EGF-like domain; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 1172
 FT DOMAIN 19 232
 FT DOMAIN 318 375
 FT DOMAIN 381 548
 FT DOMAIN 549 692
 FT DOMAIN 725 952
 FT DOMAIN 953 1172
 FT REPEAT 381 432
 FT REPEAT 437 493
 FT REPEAT 494 548

FT DOMAIN 549 EGF-LIKE 1.
FT DOMAIN 590 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 647 EGF-LIKE 3.
FT REPEAT 648 EGF-LIKE 3.
FT REPEAT 725 TSP TYPE-3 1.
FT REPEAT 761 TSP TYPE-3 2.
FT REPEAT 783 TSP TYPE-3 3.
FT REPEAT 784 TSP TYPE-3 3.
FT REPEAT 820 TSP TYPE-3 4.
FT REPEAT 843 TSP TYPE-3 5.
FT REPEAT 881 TSP TYPE-3 6.
FT REPEAT 916 TSP TYPE-3 7.
FT REPEAT 952 TSP TYPE-3 7.
FT SITE 928 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 266 INTERCHAIN (PROBABLE).
FT DISULFID 270 INTERCHAIN (PROBABLE).
FT DISULFID 270 BY SIMILARITY.
FT DISULFID 553 BY SIMILARITY.
FT DISULFID 558 BY SIMILARITY.
FT DISULFID 577 BY SIMILARITY.
FT DISULFID 594 BY SIMILARITY.
FT DISULFID 601 BY SIMILARITY.
FT DISULFID 622 BY SIMILARITY.
FT DISULFID 646 BY SIMILARITY.
FT DISULFID 652 BY SIMILARITY.
FT DISULFID 659 BY SIMILARITY.
FT DISULFID 680 BY SIMILARITY.
FT CARBOHYD 131 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 316 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 457 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 457 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 584 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 710 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1069 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1172 AA; 129955 MW; 2AC7BB230E44C6F5 CRC64;

Query Match 4.7%; Score 8; DB 1; Length 1172;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ALEPGGLS 10
Db 102 ALEPGGLS 109
|||||||

RESULT 3
IMMO_ECOLI STANDARD; PRT; 113 AA.
AC P33637;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE COLICIN EI IMMUNITY PROTEIN (IMMEL) (MICROCIN EI IMMUNITY PROTEIN).
GN IMM.
OS Escherichia coli.
OG Plasmid IncFII R100.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86319522; PubMed=3019092;
RA Ohtsubo H., Ryder T.B., Maeda Y., Armstrong K., Ohtsubo E.;
RT "DNA replication of the resistance plasmid R100 and its control.";
RL Adv. Biophys. 21:115-133(1986).
CC -!- FUNCTION: THIS PROTEIN IS ABLE TO PROTECT A CELL, WHICH HARBORS
CC THE PLASMID COLEI ENCODING COLICIN EI, AGAINST COLICIN EI.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M26840; AAA26071.1; -;
CC INTERPRO: IPR003061; -;
CC

Query Match 4.1%; Score 7; DB 1; Length 113;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 LLCLVLA 17
Db 92 LLCLVLA 98
|||||||

RESULT 4
IMM1_ECOLI STANDARD; PRT; 113 AA.
AC P02985;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE COLICIN EI IMMUNITY PROTEIN (IMMEL) (MICROCIN EI IMMUNITY PROTEIN).
GN IMM.
OS Escherichia coli.
OG Plasmid COLEI.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86094231; PubMed=3936034;
RA Waich N.S., Johnson P.H.;
RT "Structural and functional organization of the colicin EI operon.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:8389-8393(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=80010893; PubMed=384144;
RA Oka A., Nomura N., Morita M.; Sugisaki H., Sugimoto K.,
RT "Nucleotide sequence of small ColEI derivatives: structure of the
RT regions essential for autonomous replication and colicin EI
RT immunity.";
RL Mol. Gen. Genet. 172:151-159(1979).
CC -!- FUNCTION: THIS PROTEIN IS ABLE TO PROTECT A CELL, WHICH HARBORS
CC THE PLASMID COLEI ENCODING COLICIN EI, AGAINST COLICIN EI.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: J01566; AAB59122.1; -;
CC EMBL: M12543; AAA23066.1; -;
CC EMBL: V00268; CAA23529.1; -;
CC PIR: A03513; IMCECI.
CC INTERPRO: IPR003061; -;
CC PRINTS: PR01298; MICROCIN.
KW Plasmid; Bacteriocin; Colicin.
FT CONFLICT 105 105 E -> G (IN REF. 2).
SQ SEQUENCE 113 AA; 13359 MW; FFB8308B7C54E83E CRC64;

```
RESULT 5
IMM1_SHISO STANDARD; PRT; 113 AA.
AC P22558;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE COLICIN E1* IMMUNITY PROTEIN (IMMEL) (MICROCIN E1* IMMUNITY PROTEIN).
GN IMM.
OS Shigella sonnei.
OG Plasmid pKY-1.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC Shigella.
RN [1]
RP SEQUENCE FROM N.A.
RA Higashi M., Hata M., Hase T., Yamaguchi K., Masamune Y.;
RT "The nucleotide sequence of cea and the region of origin of plasmid
RT pKY-1."
RL J. Gen. Appl. Microbiol. 32:433-442(1986).
CC -!- FUNCTION: THIS PROTEIN IS ABLE TO PROTECT A CELL, WHICH HARBORS
CC THE PLASMID PKY-1 ENCODING COLICIN E1* AGAINST COLICIN E1*.
CC -!- SIMILARITY: 79% IDENTITY TO E.COLI COLICIN E1 IMMUNITY PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M37218; AAA98157.1; -
DR PIR; S11532; S11532
DR INTERPRO; IPR003061; -
DR PRINTS; PR01298; MICROCIN.
KW Plasmid; Bacteriocin; Colicin.
SQ SEQUENCE 113 AA; 13423 MW; 89A4D21E69F34155 CRC64;

Query Match 4.1%; Score 7; DB 1; Length 113;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 11 LLCLVLA 17
Db 93 LLCLVLA 99
|||||

RESULT 6
MOCA_RHIME STANDARD; PRT; 317 AA.
AC P49307;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE RHIZOPINE CATABOLISM PROTEIN MOCA.
GN MOCA.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
CC Rhizobiaceae; Sinorhizobium.
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-95147842; PubMed-7845353;
RX Rossbach S., Kuipa D.A., Rossbach U., de Bruijn F.J.;
RT "Molecular and genetic characterization of the rhizopine catabolism
RT (mocABC) genes of Rhizobium meliloti L5-30."
RL Mol. Genet. 245:11-24(1994).
CC -!- FUNCTION: COULD CATALYZE THE NADH-DEPENDENT DEHYDROGENASE REACTION
CC INVOLVED IN RHIZOPINE (L-3-O-METHYL-SCYLLO-INOSAMINE) CATABOLISM.
CC -!- SIMILARITY: BELONGS TO THE GFC/IDH/MOCA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X78503; CAA55269.1; -
DR INTERPRO; IPR000683; -
DR PFAM; PF01408; GFC_IDH_MoCA; 1.
KW Oxidoreductase; NAD.
SQ SEQUENCE 317 AA; 33725 MW; 06CD96EC77276652 CRC64;

Query Match 4.1%; Score 7; DB 1; Length 317;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 146 ARGGAPR 152
Db 277 ARGGAPR 283
|||||

RESULT 7
NOD2_BRASN STANDARD; PRT; 337 AA.
AC P50328;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE NODULATION PROTEIN D II.
GN NODD2.
OS Bradyrhizobium sp. (strain NC92).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
CC Bradyrhizobium group; Bradyrhizobium.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-96213014; PubMed-8631662;
RA Gillette W.K., Elkan G.H.;
RT "Bradyrhizobium (Arachis) sp. strain NC92 contains two nodD genes
RT involved in the repression of noda and a nola gene required for the
RT efficient nodulation of host plants."
RL J. Bacteriol. 178:2757-2766(1996).
CC -!- FUNCTION: NODD REGULATES THE EXPRESSION OF THE NOD ABCFE GENES
CC WHICH ENCODE OTHER NODULATION PROTEINS. NODD IS ALSO A NEGATIVE
CC REGULATOR OF ITS OWN EXPRESSION. BINDS FLAVONOIDS AS INDUCERS.
CC -!- MISCELLANEOUS: THERE ARE AT LEAST TWO NODD GENES IN B.SP STRAIN
CC NC92.
CC -!- SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U33192; AAB06563.1; -
DR INTERPRO; IPR000847; -
DR PFAM; PF00126; HTH_L1; 1.
DR PFAM; PF01046; NodD_C-term; 1.
DR PRINTS; PR00039; HTHLYSR.
DR PROSITE; PS00044; HTH_LYSR_FAMILY; 1.
KW Repressor; Transcription regulation; DNA-binding; Activator;
FT DNA_BIND 23 42
SQ SEQUENCE 337 AA; 38141 MW; 7645884AB5160177 CRC64;

Query Match 4.1%; Score 7; DB 1; Length 337;
```

```
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 GERLVC4 50
DB 167 GERLVC4 173

RESULT 8
CKRA_HUMAN STANDARD; PRT; 362 AA.
ID CKRA_HUMAN STANDARD; PRT; 362 AA.
AC P46092;
DT 01-NOV-1995 (Rel. 32, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE C-C CHEMOKINE RECEPTOR TYPE 10 (C-C CKR-10) (CCR-10).
GN CCR10 OR GPR2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
[1]
SEQUENCE FROM N.A., AND CHARACTERIZATION.
RP MEDLINE-20191997; PubMed-10725696;
RA Jarmin D.I., Rits M., Bota D., Gerard N.P., Graham G.J.,
RA Clark-Lewis I., Gerard C.;
RT "Cutting edge: identification of the orphan receptor G-protein-coupled
RT receptor 2 as CCR10, a specific receptor for the chemokine ESKine.";
RL J. Immunol. 164:3460-3464(2000).
[2]
SEQUENCE OF 9-362 FROM N.A.
RP MEDLINE-20191998; PubMed-10725697;
RX Honey B., Wang W., Soto H., Buchanan M.E., Wiesenborn A., Catron D.,
RA Muller A., McGlathlan T.K., Dieu-Nosjean M.C., Orozco R., Ruzicka T.,
RA Lehmann P., Oldham E., Zlotnik A.;
RT "Cutting edge: the orphan chemokine receptor G protein-coupled
RT receptor-2 (GPR-2, CCR10) binds the skin-associated chemokine CCL27
RT (CTACK/ALP/ITLC).";
RL J. Immunol. 164:3465-3470(2000).
[3]
FUNCTION: RECEPTOR FOR CHEMOKINES SCYA27 AND SCYA28. SUBSEQUENTLY
TRANSDUCE A SIGNAL BY INCREASING THE INTRACELLULAR CALCIUM IONS
LEVEL.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: EXPRESSED BY MELANOCYTES, DERMAL FIBROBLASTS,
CC AND DERMAL MICROVASCULAR ENDOTHELIAL CELLS. ALSO DETECTED IN T
CC CELLS AS WELL AS IN SKIN-DERIVED LANGERHANS CELLS.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
EMBL; AF215981; AAF63709.1; -
EMBL; U13667; AAA64593.1; -
MIM; 600240; -
DR INTERPRO; IPR000276; -
DR PFAM; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEPTOR_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEPTOR_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 52 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 53 68 1 (POTENTIAL).
```

```
FT DOMAIN 69 78 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 79 99 2 (POTENTIAL).
FT DOMAIN 100 114 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 115 136 3 (POTENTIAL).
FT DOMAIN 137 159 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 160 179 4 (POTENTIAL).
FT DOMAIN 180 203 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 204 225 5 (POTENTIAL).
FT DOMAIN 226 247 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 248 269 6 (POTENTIAL).
FT DOMAIN 270 290 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 291 313 7 (POTENTIAL).
FT DOMAIN 314 362 CYTOPLASMIC (POTENTIAL).
FT DISULFID 113 191 BY SIMILARITY.
FT CONFLICT 51 V -> L (IN REF. 2).
SQ SEQUENCE 362 AA; 38399 MW; 69DF12B639AEA99A CRC64;

Query Match 4.1%; Score 7; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LALPALL 22
DB 172 LALPALL 178

RESULT 9
PEL_LILLO STANDARD; PRT; 434 AA.
ID PEL_LILLO STANDARD; PRT; 434 AA.
AC P40973;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE PEPTATE LYASE PRECURSOR (EC 4.2.2.2).
OS Lilium longiflorum (Trumpet lily).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Liliales; Liliaceae; Lilium.
[1]
SEQUENCE FROM N.A.
RC STRAIN-CV. NELLIE WHITE; TISSUE-POLLEN;
RA Kim S.R., Finkel D.J., An G.;
RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ELIMINATIVE CLEAVAGE OF PEPTATE TO GIVE
CC OLIGOSACCHARIDES WITH 4-DEOXY-ALPHA-D-MANN-4-ENURONOSYL GROUPS
CC AT THEIR NON-REDUCING ENDS.
CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
EMBL; Z17328; CAA78976.1; -
EMBL; L18911; AAA33398.1; -
DR PIR; S29612; S29612.
DR INTERPRO; IPR002022; -
DR PFAM; PF00544; pec_lyase; 1.
DR PRINTS; PR00807; AMBALLERGEN.
KW Lyase; Signal.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 434 PEPTATE LYASE.
FT ACT_SITE 312 312 POTENTIAL.
FT CARBOHYD 68 68 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 97 97 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 434 AA; 48457 MW; C1F3E30AD2BBD064 CRC64;

Query Match 4.1%; Score 7; DB 1; Length 434;
Best Local Similarity 100.0%; Pred. No. 22;
```

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 PGTFVQR 60
 Db 412 PGTFVQR 418
 |||||

RESULT 10
 RRB_MOUSE
 ID RRB_MOUSE STANDARD; PRT; 482 AA.
 AC P22605; P22604; P11417;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE RETINOIC ACID RECEPTOR BETA (RAR-BETA)
 GN RARB OR NR1B2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN [1]
 RP SEQUENCE FROM N.A. (RAR-BETA-1; RAR-BETA-2 AND RAR-BETA-3).
 RC STRAIN=C57BL/C;
 RX MEDLINE=91114714; PubMed=1846599;
 RA Zelent A., Mendelsohn C., Kastner P., Krust A., Garnier J.-M.,
 RA Ruffenach F., Leroy P., Chambon P.;
 RT "Differentially expressed isoforms of the mouse retinoic acid
 RT receptor beta generated by usage of two promoters and alternative
 RT splicing.";
 RL EMBO J. 10:71-81(1991).
 RN [2]
 RP SEQUENCE FROM N.A. (RAR-BETA-2).
 RX MEDLINE=9295563; PubMed=2544807;
 RA Zelent A., Krust A., Petkovitch M., Kastner P., Chambon P.;
 RT "Cloning of murine alpha and beta retinoic acid receptors and a novel
 RT receptor gamma predominantly expressed in skin.";
 RL Nature 339:714-717(1989).
 RN [3]
 RP SEQUENCE FROM N.A. (RAR-BETA-2).
 RX MEDLINE=93195858; PubMed=8383767;
 RA Heilmann R., Rentrop M., Lang E., Maelicke A.;
 RT "Cloning of several genes coding for retinoic acid nuclear receptors
 RT in the mouse embryonal carcinoma cell line PCC7-MZ1.";
 RL J. Recept. Res. 13:693-709(1993).
 RN [4]
 RP SEQUENCE FROM N.A. (RAR-BETA-4).
 RX MEDLINE=92212900; PubMed=1313565;
 RA Nagpal S., Zelent A., Chambon P.;
 RT "RAR-beta 4, a retinoic acid receptor isoform is generated from
 RT RAR-beta 2 by alternative splicing and usage of a CUG initiator
 RT codon.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2718-2722(1992).
 CC -1- FUNCTION: THIS IS A RECEPTOR FOR RETINOIC ACID. THIS METABOLITE
 CC HAS PROFOUND EFFECTS ON VERTEBRATE DEVELOPMENT. RETINOIC ACID IS
 CC A MORPHOGEN AND IS A POWERFUL TERATOGEN. THIS RECEPTOR CONTROLS
 CC CELLS FUNCTIONS BY DIRECTLY REGULATING GENE EXPRESSION.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS; RAR-BETA-1, RAR-BETA-2, RAR-
 CC BETA-3 (SHOWN HERE) AND RAR-BETA-4; ARE PRODUCED BY ALTERNATIVE
 CC SPLICING.
 CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
 CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 CC NRI SUBFAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL; X56574; CAA39921.1; "

DR EMBL; X56569; CAA39918.1; "
 DR EMBL; X56573; CAA39920.1; "
 DR EMBL; S56600; CAB31483.1; "
 DR EMBL; S92180; "; NOT_ANNOTATED_CDS.
 DR PIR; S13515; S13515.
 DR PIR; S13516; S13516.
 DR PIR; S13517; S13517.
 DR PIR; S05051; S05051.
 DR HSP; P10826; IHRA.
 DR TRANSFAC; T00717; "
 DR TRANSFAC; T01339; "
 DR MGD; MGI:97857; RARB.
 DR INTERPRO; IPR000536; "
 DR INTERPRO; IPR001628; "
 DR INTERPRO; IPR001723; "
 DR INTERPRO; IPR003078; "
 DR PFAM; PF00104; hormone_rec.1.
 DR PFAM; PF00105; zf-C4; 1.
 DR PRINTS; PR00047; STRDIDFINGER.
 DR PRINTS; PR00398; STRDHORMONER.
 DR PRINTS; PR01292; RETNOICACIDR.
 DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
 KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
 KW Zinc-finger; Multigene family; Alternative splicing.
 FT DOMAIN 1 114
 FT DNA_BIND 115 180 C4-TYPE ZINC FINGERS (TWO).
 FT ZN_FING 115 135 C4-TYPE.
 FT ZN_FING 151 175 C4-TYPE.
 FT DOMAIN 181 226 HINGE.
 FT DOMAIN 227 446 LIGAND-BINDING.
 FT VARSPLIC 1 87
 FT VARSPLIC 1 83 MISSING (IN ISOFORM RAR-BETA-4).
 FT VARSPLIC 84 86 LNA -> MEN (IN ISOFORM RAR-BETA-4).
 FT VARSPLIC 60 60 S -> T (IN ISOFORM RAR-BETA-1).
 FT VARSPLIC 61 87 MISSING (IN ISOFORM RAR-BETA-1).
 SQ SEQUENCE 482 AA; 53331 MW; 48E78E6C7D012515 CRC64;
 MSTS SHACPPAVRGHMTHTYPAAPYPLLPVIRGLSLPPL
 HGLHGHPPSGCSTPSPASVQACQRTTGGQFAASTKWT
 SLNAA -> MEDCMDVLSVSGOILDEVTASPSMLOEKA
 LKACLSGFTQAEWHRTAQS (IN ISOFORM RAR-
 BETA-2).
 Query Match 4.1%; Score 7; DB 1; Length 482;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 PVPVARG 29
 |||||
 Db 9 PVPVARG 15

RESULT 11
 TTD_TCOLI
 ID TTD_TCOLI STANDARD; PRT; 487 AA.
 AC P39414; O46870;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE PUTATIVE TARTRATE CARRIER (TARTRATE TRANSPORTER) (TARTRATE/SUCCINATE
 DE ANTI-PORTER).
 GN YGJE.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AF100318; AAD05304.1; -;
 DR MIM; 604468; -;
 DR INTERPRO; IPR000719; -;
 DR INTERPRO; IPR001245; -;
 DR PFAM; PF00069; pkinase; 1;
 DR PRINTS; PR00109; TYRKINASE.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding.
 FT DOMAIN 371 629 PROTEIN KINASE
 FT NP_BIND 377 384 ATP (BY SIMILARITY).
 FT BINDING 400 400 ATP (BY SIMILARITY).
 FT ACT_SITE 494 494 BY SIMILARITY.
 SQ SEQUENCE 1011 AA; 112512 MW; F7294BA0587D2EC9 CRC64;

Query Match 4.1%; Score 7; DB 1; Length 1011;

Best Local Similarity 100.0%; Pred. No. 44;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 ETGERLV 48

Db 370 ETGERLV 376

RESULT 16

APB_HUMAN STANDARD; PRT; 4563 AA.
 ID P04114; O00502; Q13787;
 AC 01-NOV-1986 (Rel. 03, Created)
 DT 01-NOV-1986 (Rel. 03, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE APOLIPOPROTEIN B-100 PRECURSOR (APO B-100) [CONTAINS: APOLIPOPROTEIN
 DE B-48 (APO B-48)].
 GN APOB.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP MEDLINE=87016385; PubMed=3763409;
 RA Knott T.C., Wallis S.C., Powell L.M., Pease R.J., Lusis A.J.;
 RA Blackhart B., McCarthy B.J., Mahley R.W., Levy-Wilson B., Scott J.;
 RT "Complete cDNA and derived protein sequence of human apolipoprotein
 RT B-100.";
 RL Nucleic Acids Res. 14:7501-7503(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88003974; PubMed=3652907;
 RA Ludwig E.H., Blackhart B.D., Pierotti V.R., Caiati L., Fortier C.,
 RA Knott T., Scott J., Mahley R.W., Levy-Wilson B., McCarthy B.J.;
 RT "DNA sequence of the human apolipoprotein B gene.";
 RL DNA 6:363-372(1987).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87008488; PubMed=3759943;
 RA Chen S.-H., Yang C.-Y., Chen P.-F., Setzer D., Tanimura M., Li W.-H.,
 RA Gotto A.M. Jr., Chan L.;
 RT "The complete cDNA and amino acid sequence of human apolipoprotein
 RT B-100.";
 RL J. Biol. Chem. 261:12918-12921(1986).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87041416; PubMed=3464946;
 RA Law S.W., Grant S.M., Higuchi K., Hospattankar A.V., Lackner K.J.,
 RA Lee N., Brewer H.B. Jr.;
 RT "Human liver apolipoprotein B-100 cDNA: complete nucleic acid and
 RT derived amino acid sequence.";

RL Proc. Natl. Acad. Sci. U.S.A. 83:8142-8146(1986).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87161758; PubMed=3030729;
 RA Cladaras C., Hadzopoulou-Cladaras M., Nolte R.T., Atkinson D.,
 RA Zannis V.I.;
 RT "The complete sequence and structural analysis of human
 RT apolipoprotein B-100: relationship between apoB-100 and apoB-48
 RT forms.";
 RL EMBO J. 5:3495-3507(1986).
 RN [6]
 RP SEQUENCE OF 709-906 FROM N.A.
 RX MEDLINE=85270450; PubMed=3860836;
 RA Deeb S.S., Motulsky A.G., Albers J.J.;
 RT "A partial cDNA clone for human apolipoprotein B.";
 RT Proc. Natl. Acad. Sci. U.S.A. 82:4983-4986(1985).
 RN [7]
 RP SEQUENCE OF 3056-3159 FROM N.A.
 RX MEDLINE=86041888; PubMed=3903660;
 RA Mehrabian M., Schumaker V.N., Fareed G.C., West R., Johnson D.F.,
 RA Kirchgessner T.G., Lin H.-C., Wang X., Ma Y., Mendiaz E., Lusis A.J.;
 RT "Human apolipoprotein B: identification of cDNA clones and
 RT characterization of mRNA.";
 RL Nucleic Acids Res. 13:6937-6953(1985).
 RN [8]
 RP SEQUENCE OF 1937-2018 AND 3811-4334 FROM N.A.
 RX MEDLINE=86093680; PubMed=3841204;
 RA Carlsson P., Olofsson S.O., Bondjers G., Wiklund O.,
 RA Bjursell G.;
 RT "Molecular cloning of human apolipoprotein B cDNA.";
 RL Nucleic Acids Res. 13:8813-8826(1985).
 RN [9]
 RP SEQUENCE OF 3109-4563 FROM N.A.
 RX MEDLINE=85300528; PubMed=2994225;
 RA Knott T.J., Rall S.C. Jr., Innerarity T.L., Jacobson S.F.,
 RA Uredea M.S., Levy-Wilson B., Powell L.M., Pease R.J., Eddy R.,
 RA Nakai H., Byers M., Priestley L.M., Robertson E., Rall L.B.,
 RA Beisholtz C., Shows T.B., Mahley R.W., Scott J.;
 RT "Human apolipoprotein B: structure of carboxyl-terminal domains,
 RT sites of gene expression, and chromosomal localization.";
 RL Science 230:37-43(1985).
 RN [10]
 RP SEQUENCE OF 1-291 FROM N.A.
 RX MEDLINE=86149525; PubMed=3513177;
 RA Protter A.A., Hardman D.A., Schilling J.W., Miller J., Appleby V.,
 RA Chen G.C., Kirsher S.W., McEnroe G., Kane J.P.;
 RT "Isolation of a cDNA clone encoding the amino-terminal region of
 RT human apolipoprotein B.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:1467-1471(1986).
 RN [11]
 RP SEQUENCE OF 1-1670 FROM N.A.
 RX MEDLINE=86287319; PubMed=3461454;
 RA Protter A.A., Hardman D.A., Sato K.Y., Schilling J.W.,
 RA Yamanaka M., Hort Y.J., Hjerrild K.A., Chen G.C., Kane J.P.;
 RT "Analysis of cDNA clones encoding the entire B-26 region of human
 RT apolipoprotein B.";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:5678-5682(1986).
 RN [12]
 RP PARTIAL SEQUENCE, AND IDENTIFICATION (APO-B48).
 RX MEDLINE=88018019; PubMed=3659919;
 RA Chen S.-H., Habib G., Yang C.-H., Gu Z.-W., Lee B.R., Weng S.-H.,
 RA Silberman S.R., Cai S.-J., Deslypere J.P., Rosseneu M.,
 RA Gotto A.M. Jr., Li W.-H., Chan L.;
 RT "Apolipoprotein B-48 is the product of a messenger RNA with an organ-
 RT specific in-frame stop codon.";
 RL Science 238:363-366(1987).
 RN [13]
 RP DOMAINS.
 RX MEDLINE=87039351; PubMed=3773997;
 RA Knott T.C., Pease R.J., Powell L.M., Wallis S.C., Rall S.C. Jr.,
 RA Innerarity T.L., Blackhart B., Taylor W.R., Marcel Y., Milne R.,
 RA Johnson D., Fuller M., Lusis A.J., McCarthy B.J., Mahley R.W.,
 RA Levy-Wilson B., Scott J.;

Complete protein sequence and identification of structural domains of human apolipoprotein B.;
Nature 323:734-738(1986).
[14]

RA Yang C.-Y., Chen S.-H., Gianturco S.H., Bradley W.A., Sparrow J.T.,
Tanimura M., Li W.-H., Sparrow D.A., Deloof H., Rosseneu M.,
Lee F.-S., Gu Z.-W., Gotto A.M. Jr., Chan L.;
"Sequence, structure, receptor-binding domains and internal repeats
of human apolipoprotein B-100.";
Nature 323:738-742(1986).
[15]

RA CALCIUM-BINDING DATA.
MEDLINE-86242245; PubMed=3087360;
Dashti N., Lee D.M., Mok T.;
"Apolipoprotein B is a calcium binding protein.";
Biochem. Biophys. Res. Commun. 137:493-499(1986).
[16]

RA VARIANT SER-4338.
MEDLINE-91071750; PubMed=1979313;
Navares M., Laurent A.-M., Moreel J.-F., Ragab A., Cambou J.-P.,
Cunha G., Cambien F., Roizes G.;
"Detection by denaturing gradient gel electrophoresis of a new
polymorphism in the apolipoprotein B gene.";
Hum. Genet. 86:91-93(1990).
[17]

RA VARIANT FDB GLN-3527.
MEDLINE-89089975; PubMed=2563166;
Soria L.F., Ludwig E.H., Clarke H.R.G., Vega G.L., Grundy S.M.,
McCarthy B.J.;
"Association between a specific apolipoprotein B mutation and
familial defective apolipoprotein B-100.";
Proc. Natl. Acad. Sci. U.S.A. 86:587-591(1989).
[18]

RA VARIANT LEU-2739.
MEDLINE-91016974; PubMed=2216805;
Huang L.-S., Gavish D., Breslow J.L.;
"Sequence polymorphism in the human apob gene at position 8344.";
Nucleic Acids Res. 18:5922-5922(1990).
[19]

RA VARIANT FDB CYS-3558.
MEDLINE-95190020; PubMed=7883971;
Pullinger C.R., Hennessy L.K., Chatterton J.E., Liu W., Love J.A.,
Mendel C.M., Frost P.H., Malloy M.J., Schumaker V.N., Kane J.P.;
"Familial ligand-defective apolipoprotein B. Identification of a new
mutation that decreases LDL receptor binding affinity.";
J. Clin. Invest. 95:1225-1234(1995).
[20]

RA VARIANTS L-1437; S-1914; K-2566; T-3121; A-3945; M-4128 AND T-4481.
MEDLINE-97044521; PubMed=8889592;
Poirier O., Ricard S., Bebaque I., Souriau C., Evans A.E.,
Arveiler D., Marques-Vidal P., Luc G., Roizes G., Cambien F.;
"Detection of new variants in the apolipoprotein B (Apo B) gene by
PCR-SSCP.";
Hum. Mutat. 8:282-285(1996).
[21]

RA VARIANTS FDB GLN-3527 AND CYS-3558.
MEDLINE-97403938; PubMed=9259199;
Rabes J.P., Varret M., Saint-Jore B., Erlich D., Jondeau G.,
Kremf M., Giraudet P., Junien C., Boileau C.;
"Familial ligand-defective apolipoprotein B-100: simultaneous
detection of the ARG3500-->GLN and ARG3531-->CYS mutations in a
French population.";
Hum. Mutat. 10:160-163(1997).
[22]

RA VARIANTS S-1914; R-1923; L-2739; D-3319; T-3427; Q-3432 AND I-3921.
MEDLINE-98141125; PubMed=9490296;
Leren T.P., Bakken K.S., Hoel V., Hjermann I., Berg K.;
"Screening for mutations of the apolipoprotein B gene causing
hypocholesterolemia.";
Hum. Genet. 102:44-49(1998).
[23]

CC -!- FUNCTION: APOLOPROTEIN B IS A MAJOR PROTEIN CONSTITUENT OF
CHYLOMICRONS, VLDL AND LDL. IT FUNCTIONS AS A RECOGNITION SIGNAL

CC FOR THE CELLULAR BINDING AND INTERNALIZATION OF LDL PARTICLES BY
CC THE APOB/E RECEPTOR.
CC -!- DISEASE: DEFECTS IN APOB ARE A CAUSE OF FAMILIAL LIGAND-DEFECTIVE
CC OF LIPOPROTEIN METABOLISM LEADING TO HYPERCHOLESTEROLEMIA AND
CC INCREASED PRONENESS TO CORONARY ARTERY DISEASE (CAD).
CC THE PLASMA CHOLESTEROL LEVELS ARE DRAMATICALLY ELEVATED DUE TO
CC IMPAIRED CLEARANCE OF LDL PARTICLES BY DEFECTIVE APOB/E RECEPTORS.
CC -!- DISEASE: DEFECTS IN APOB ASSOCIATED WITH OTHER GENE DEFECTS
CC (POLYGENIC) CAN BE THE CAUSE OF HYPOCHOLESTEROLEMIA.
CC -!- MISCELLANEOUS: APO B-48 WHICH IS SYNTHESIZED ONLY BY THE
CC INTESTINE, AND FOUND IN CHYLOMICRONS, IS A SHORTENED FORM OF APO
CC B-100 WHICH LACK THE LDL-RECEPTOR REGION. B-48 IS PRODUCED BY THE

Query Match 4.18; Score 7; DB 1; Length 4563;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LALPALL 22

DB 12 LALPALL 18

|||||

RESULT 17

XYLB ACTMI

ID XYLB.ACTMI STANDARD; PRT; 61 AA.

AC P12867;

DT 01-OCT-1989 (Rel. 12, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 01-FEB-1996 (Rel. 33, Last annotation update)

DE XYLUOSE KINASE (EC 2.7.1.17) (XYLUOKINASE) (FRAGMENT).

GN XYLB.

OS Actinoplanes missouriensis.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Micromonosporineae; Micromonosporaceae;

OC Actinoplanes.

RN [1]

RC SEQUENCE FROM N.A.

RP STRAIN-DSM 43046;

RP MEDLINE-90016811; PubMed=2798103;

RA Amore R., Hollenberg C.P.;

RT "Xylose isomerase from Actinoplanes missouriensis: primary structure

of the gene and the protein.";

RL Nucleic Acids Res. 17:7515-7515(1989).

CC -!- CATALYTIC ACTIVITY: ATP + D-XYLUOSE -> ADP + D-XYLUOSE

CC 5-PHOSPHATE.

CC -!- SIMILARITY: BELONGS TO THE FUCOKINASE / GLUCOKINASE /

CC GLYCEROKINASE / XYLUOKINASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

CC EMBL; X15042; CAA34165.1; -

CC PIR; S05999; S05999.

CC INTERPRO; IPR000577; -

DR PROSITE; PS00445; FGGY_KINASES_2; PARTIAL.

DR PROSITE; PS00933; FGGY_KINASES_1; PARTIAL.

KW Transferase; Kinase; Xylose metabolism.

FT NON_TER 61

SQ SEQUENCE 61 AA; 6423 MW; D1B6806BC66C1695 CRC64;

Query Match 3.58; Score 6; DB 1; Length 61;

Best Local Similarity 100.0%; Pred. No. 46;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 RDAETG 44

|||||

```
Db 19 RDAETG 24

RESULT 18
VG30_BPM2D STANDARD; PRT; 88 AA.
AC O64224;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GENE 30 PROTEIN (GP30).
GN 30.
OS Mycobacteriophage D29.
OC Viruses.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98300335; PubMed=9636706;
RA Ford M.E., Sarkis G.J., Belanger A.E., Hendrix R.W., Hatfull G.F.;
RT "Genome structure of mycobacteriophage D29: implications for phage
evolution.";
RL J. Mol. Biol. 279:143-164(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF022214; AAC18471.1;
SQ SEQUENCE 88 AA; 10391 MW; 52071A2E676E5BCE CRC64;

Query Match 3.5%; Score 6; DB 1; Length 88;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 LPALLP 23
Db 25 LPALLP 30
|||||

RESULT 19
HRK_MOUSE STANDARD; PRT; 92 AA.
AC P70678;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE ACTIVATOR OF APOPTOSIS HARAKIRI (NEURONAL DEATH PROTEIN DP5).
GN HRK OR DP5 OR BID3.
OS Mus musculus (Mouse), and Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=BRAIN;
RA Inaizumi K.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ACTIVATES APOPTOSIS AND INTERACTS SELECTIVELY WITH
CC SURVIVAL-PROMOTING PROTEINS BCL-2 AND BCL-XL (BY SIMILARITY).
CC -!- DOMAIN: THE BH3 DOMAIN IS REQUIRED FOR THE INDUCTION OF CELL
CC DEATH (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS A BCL-2 HOMOLGY DOMAIN 3 (BH3).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF022214; AAC18471.1;
SQ SEQUENCE 88 AA; 10391 MW; 52071A2E676E5BCE CRC64;

Db 19 RDAETG 24

RESULT 18
VG30_BPM2D STANDARD; PRT; 88 AA.
AC O64224;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GENE 30 PROTEIN (GP30).
GN 30.
OS Mycobacteriophage D29.
OC Viruses.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98300335; PubMed=9636706;
RA Ford M.E., Sarkis G.J., Belanger A.E., Hendrix R.W., Hatfull G.F.;
RT "Genome structure of mycobacteriophage D29: implications for phage
evolution.";
RL J. Mol. Biol. 279:143-164(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF022214; AAC18471.1;
SQ SEQUENCE 88 AA; 10391 MW; 52071A2E676E5BCE CRC64;

Query Match 3.5%; Score 6; DB 1; Length 88;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 LPALLP 23
Db 25 LPALLP 30
|||||

RESULT 19
HRK_MOUSE STANDARD; PRT; 92 AA.
AC P70678;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE ACTIVATOR OF APOPTOSIS HARAKIRI (NEURONAL DEATH PROTEIN DP5).
GN HRK OR DP5 OR BID3.
OS Mus musculus (Mouse), and Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=BRAIN;
RA Inaizumi K.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ACTIVATES APOPTOSIS AND INTERACTS SELECTIVELY WITH
CC SURVIVAL-PROMOTING PROTEINS BCL-2 AND BCL-XL (BY SIMILARITY).
CC -!- DOMAIN: THE BH3 DOMAIN IS REQUIRED FOR THE INDUCTION OF CELL
CC DEATH (BY SIMILARITY).
CC -!- SIMILARITY: CONTAINS A BCL-2 HOMOLGY DOMAIN 3 (BH3).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF022214; AAC18471.1;
SQ SEQUENCE 88 AA; 10391 MW; 52071A2E676E5BCE CRC64;

Query Match 3.5%; Score 6; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 LPALLP 23
Db 59 LPALLP 64
|||||

RESULT 20
LE10_HELAN STANDARD; PRT; 92 AA.
AC P46514;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 10 KDA LATE EMBRYOGENESIS ABUNDANT PROTEIN (DS10).
OS Helianthus annuus (Common sunflower).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae;
OC euasterids II; Asterales; Asteraceae; Asteroideae; Heliantheae;
OC Helianthus.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CV. SUNWEED; TISSUE=DRY SEED;
RX MEDLINE=92353387; PubMed=1386536;
RA Almoquera C., Jordano J.;
RT "Developmental and environmental concurrent expression of sunflower
RT dry-seed-stored low-molecular-weight heat-shock protein and Lea
RT mRNAs.";
RL Plant Mol. Biol. 19:781-792(1992).
RN [2]
RP REVISION TO 51.
RX STRAIN=CV. SUNWEED; TISSUE=COTYLEDON;
RA Jordano J.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=99190608; PubMed=10092187;
RA Prieto-Dapena P., Almoquera C., Rojas A., Jordano J.;
RT "Seed-specific expression patterns and regulation by ABI3 of an
RT unusual late embryogenesis abundant gene in sunflower.";
RL Plant Mol. Biol. 39:615-627(1999).
CC -!- FUNCTION: LEA PROTEINS ARE LATE EMBRYONIC PROTEINS ABUNDANT IN
CC HIGHER PLANT SEED EMBRYOS. THEY MAY PLAY AN ESSENTIAL ROLE IN
CC SEED SURVIVAL AND IN CONTROLLING WATER EXCHANGES DURING SEED
CC DESICCATION AND IMBIBITION.
CC -!- TISSUE SPECIFICITY: MAXIMALLY EXPRESSED IN DRY SEEDS. ALSO
CC PRESENT IN MID-MATURATION EMBRYOS.
CC -!- INDUCTION: BY ABSCISIC ACID, OSMOTIC STRESS AND HEAT SHOCK.
CC -!- SIMILARITY: BELONGS TO THE SMALL HYDROPHILIC PLANT SEED PROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X59699; CAA42220.1;
SQ SEQUENCE 92 AA; 10078 MW; 95760B9EAE88F1DE CRC64;

Query Match 3.5%; Score 6; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 LPALLP 23
Db 59 LPALLP 64
|||||

RESULT 20
LE10_HELAN STANDARD; PRT; 92 AA.
AC P46514;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 10 KDA LATE EMBRYOGENESIS ABUNDANT PROTEIN (DS10).
OS Helianthus annuus (Common sunflower).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae;
OC euasterids II; Asterales; Asteraceae; Asteroideae; Heliantheae;
OC Helianthus.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CV. SUNWEED; TISSUE=DRY SEED;
RX MEDLINE=92353387; PubMed=1386536;
RA Almoquera C., Jordano J.;
RT "Developmental and environmental concurrent expression of sunflower
RT dry-seed-stored low-molecular-weight heat-shock protein and Lea
RT mRNAs.";
RL Plant Mol. Biol. 19:781-792(1992).
RN [2]
RP REVISION TO 51.
RX STRAIN=CV. SUNWEED; TISSUE=COTYLEDON;
RA Jordano J.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=99190608; PubMed=10092187;
RA Prieto-Dapena P., Almoquera C., Rojas A., Jordano J.;
RT "Seed-specific expression patterns and regulation by ABI3 of an
RT unusual late embryogenesis abundant gene in sunflower.";
RL Plant Mol. Biol. 39:615-627(1999).
CC -!- FUNCTION: LEA PROTEINS ARE LATE EMBRYONIC PROTEINS ABUNDANT IN
CC HIGHER PLANT SEED EMBRYOS. THEY MAY PLAY AN ESSENTIAL ROLE IN
CC SEED SURVIVAL AND IN CONTROLLING WATER EXCHANGES DURING SEED
CC DESICCATION AND IMBIBITION.
CC -!- TISSUE SPECIFICITY: MAXIMALLY EXPRESSED IN DRY SEEDS. ALSO
CC PRESENT IN MID-MATURATION EMBRYOS.
CC -!- INDUCTION: BY ABSCISIC ACID, OSMOTIC STRESS AND HEAT SHOCK.
CC -!- SIMILARITY: BELONGS TO THE SMALL HYDROPHILIC PLANT SEED PROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X59699; CAA42220.1;
SQ SEQUENCE 92 AA; 10078 MW; 95760B9EAE88F1DE CRC64;
```

```

DR EMBL; AJ224116; CAA11834.1; -.
DR INTERPRO; IPR000389; -.
DR PFAM; PF00477; seed.protein; 1.
DR PROSITE; PS00431; SMALL_HYDR_PLANT_SEED; 1.
KW Seed embryo; Seed.
SQ SEQUENCE 92 AA; 10037 MW; BD612FA4CE72BAFF CRC64;

Query Match          3.5%; Score 6; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 GEREE 101
    |||||
DB 83 GEREE 88

RESULT 21
FIXX_RHILP
ID FIXX_RHILP STANDARD; PRT; 98 AA.
AC P08710;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE FERREDOXIN LIKE PROTEIN.
FIXX
OS Rhizobium leguminosarum (biovar trifolii).
OG Plasmid sym.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ANU 843;
RX MEDLINE=87174837; PubMed=3562251;
RA Iismaa S.E., Watson J.M.;
RT "A gene upstream of the Rhizobium trifolii nifA gene encodes a
RT ferredoxin-like protein."
RL Nucleic Acids Res. 15:3180-3180(1987).
CC -!- FUNCTION: COULD BE A 3FE-4S CLUSTER-CONTAINING PROTEIN.
CC -!- SIMILARITY: TO FERREDOXINS FROM PSEUDOMONAS PUTIDA AND CLOSTRIDIUM
CC TARTARIVORUM, FERREDOXIN I FROM ACETOBACTER VINELANDII, FERREDOXIN
CC II FROM DESULFOVIBRIO DESULFURICANS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X05257; CAA28879.1; -.
DR PIR; A27510; A27510.
DR INTERPRO; IPR001450; -.
DR PROSITE; PS00198; 4FE4S_FERREDOXIN; 1.
KW Electron transport; Iron-sulfur; Nitrogen fixation; Plasmid.
SQ SEQUENCE 98 AA; 11110 MW; 6BC3047472E1F764 CRC64;

Query Match          3.5%; Score 6; DB 1; Length 98;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 CNVLGC 96
    |||||
DB 72 CNVLGC 77

RESULT 22
YXCD_BACSU
ID YXCD_BACSU STANDARD; PRT; 102 AA.
AC P46334;
DT 01-NOV-1995 (Rel. 32, Created)

```

```

DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYPOTHETICAL 11.9 KDA PROTEIN IN IOLS-HTPG INTERGENIC REGION.
GN YXCD OR SS92CR.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / BGSCIAL;
RX MEDLINE=96093926; PubMed=7584049;
RA Yoshida K.-I., Seki S., Fujimura M., Miwa Y., Fujita Y.;
RT "Cloning and sequencing of a 36-kb region of the Bacillus subtilis
RT genome between the gnt and iol operons."
RL DNA Res. 2:61-69(1995).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AB005534; BAA21605.1; -.
DR EMBL; Z99124; CAB16016.1; -.
DR SUBTILIST; BG11361; YXCD.
KW Hypothetical protein.
SQ SEQUENCE 102 AA; 11851 MW; A8691BC2EF34656C CRC64;

Query Match          3.5%; Score 6; DB 1; Length 102;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 116 RTGFFA 121
    |||||
DB 39 RTGFFA 44

RESULT 23
NOLE_RHILP
ID NOLE_RHILP STANDARD; PRT; 110 AA.
AC P23716;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-MAR-1992 (Rel. 21, Last annotation update)
DE MODULATION PROTEIN NOLE PRECURSOR.
GN NOLE.
OS Rhizobium leguminosarum (biovar phaseoli).
OG Plasmid sym.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8002;
RX MEDLINE=91014692; PubMed=2215216;
RA Davis E.O., Johnston A.W.B.;
RT "Analysis of three nodD genes in Rhizobium leguminosarum biovar
RT phaseoli; nodB1 is preceded by noIE, a gene whose product is secreted
RT from the cytoplasm."
RL Mol. Microbiol. 4:921-932(1990).
CC -!- SUBCELLULAR LOCATION: PERIPLASMIC (POTENTIAL).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X54214; CAA38125.1; -.

```

DR PIR; S11786; S11786.
 KW Plasmid; Nodulation; Periplasmic; Signal.
 FT SIGNAL 1 ? POTENTIAL.
 FT CHAIN ? 110 NODULATION PROTEIN NOLE.
 SQ SEQUENCE 110 AA; 12050 MW; EF2FB42A7BDBC904 CRC64;

Query Match 3.5%; Score 6; DB 1; Length 110;
 Best Local Similarity 100.0%; Pred. No. 73;
 Matches 6; Conservative 0; Mismatches 0; Gaps 0;

QY 164 AGPSLA 169
 |||||
 DB 20 AGPSLA 25

RESULT 24
 SC4_SCHCO STANDARD; PRT; 111 AA.
 ID SC4_SCHCO
 AC PL6934;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE FRUITING BODY PROTEIN SC4 PRECURSOR (HYDROPHOBIN SC4).
 GN SC4.
 OS Schizophyllum commune (Bracket fungus).
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Stereales;
 OC Schizophyllaceae; Schizophyllum.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90382666; PubMed=2401401;
 RA Schuren F.H.J., Wessels J.G.H.;
 RT "Two genes specifically expressed in fruiting dikaryons of
 RT Schizophyllum commune: homologues with a gene not regulated by
 RT mating-type genes";
 RL Gene 90:199-205(1990).
 CC -1- FUNCTION: CONTRIBUTES TO SURFACE HYDROPHOBICITY, WHICH IS
 CC IMPORTANT FOR PROCESSES SUCH AS ASSOCIATION OF HYPHAE IN
 CC REPRODUCTIVE STRUCTURES, DISPERSAL OF AERIAL SPORES AND ADHESION
 CC OF PATHOGENS TO HOST STRUCTURES.
 CC -1- SUBUNIT: IN SOLUTION, THE PROTEIN MONOMERS FORM RODLET-LIKE AND
 CC INSOLUBLE AGGREGATES.
 CC -1- SUBCELLULAR LOCATION: CELL WALL; ABUNDANTLY SECRETED (IN AQUEOUS
 CC ENVIRONMENT).
 CC -1- DEVELOPMENTAL STAGE: IS ABUNDANTLY EXPRESSED AND ACCUMULATES IN
 CC THE WALLS OF DEVELOPING FRUITING BODIES (ONLY IN FRUITING
 CC DIKARYONS).
 CC -1- SIMILARITY: BELONGS TO THE FUNGAL HYDROPHOBIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; M32330; AAA33927.1; ..
 DR PIR; JH0183; JH0183.
 DR INTERPRO: IPR001338; ..
 DR PFAM; PF01185; Hydrophobin; 1.
 DR PROSITE; PS00956; HYDROPHOBIN; 1.
 KW Fruiting body; Glycoprotein; Signal.
 FT SIGNAL 1 25 POTENTIAL.
 FT CHAIN 26 111 FRUITING BODY PROTEIN SC4.
 FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 71 71 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 111 AA; 10729 MW; FEBEB2B2CAR846DF CRC64;

Query Match 3.5%; Score 6; DB 1; Length 111;
 Best Local Similarity 100.0%; Pred. No. 74;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LALPAL 21
 |||||
 DB 8 LALPAL 13

RESULT 25
 NEUL_HUMAN STANDARD; PRT; 125 AA.
 ID NEUL_HUMAN
 AC P01178;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE OXYTOCIN-NEUROPHYSIN 1 PRECURSOR (OT-NPT) [CONTAINS: OXYTOCIN
 DE (OCYTOCIN); NEUROPHYSIN 1].
 GN OXT OR OT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87026235; PubMed=3768139;
 RA Rehbein M., Hillers M., Mohr E., Ivell R., Morley S., Schmale H.,
 RA Richter D.;
 RT "The neurohypophyseal hormones vasopressin and oxytocin. Precursor
 RT structure, synthesis and regulation."
 RL Biol. Chem. Hoppe-Seyler 367:695-704(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85261445; PubMed=2991279;
 RA Sausville E., Carney D., Battey J.;
 RT "The human vasopressin gene is linked to the oxytocin gene and is
 RT selectively expressed in a cultured lung cancer cell line."
 RL J. Biol. Chem. 260:10236-10241(1985).
 RN [3]
 RP SEQUENCE OF 25-110 FROM N.A.
 RX MEDLINE=91065270; PubMed=2249637;
 RA Ivell R., Furiya K., Brackmann B., Dawood Y., Khan-Dawood F.;
 RT "Expression of the oxytocin and vasopressin genes in human and baboon
 RT gonadal tissues."
 RL Endocrinology 127:2990-2996(1990).
 RN [4]
 RP SEQUENCE OF 80-125 FROM N.A.
 RX MEDLINE=86056283; PubMed=4065330;
 RA Mohr E., Hillers M., Ivell R., Haulica I.D., Richter D.;
 RT "Expression of the vasopressin and oxytocin genes in human
 RT hypothalamus."
 RL FEBS Lett. 193:12-16(1985).
 RN [5]
 RP SEQUENCE OF 32-125.
 RX MEDLINE=83221497; PubMed=6574452;
 RA Chauvet M.-T., Hurpet D., Chauvet J., Acher R.;
 RT "Identification of human neurophysins: complete amino acid sequences
 RT of MSEL- and VLDV-neurophysins."
 RL Proc. Natl. Acad. Sci. U.S.A. 80:2839-2843(1983).
 RN [6]
 RP SEQUENCE OF 32-125.
 RX MEDLINE=81261248; PubMed=7262323;
 RA Schlesinger D.H., Audhya T.K.;
 RT "A comparative study of mammalian neurophysin protein sequences."
 RL FEBS Lett. 128:325-328(1981).
 RN [7]
 RP SEQUENCE OF 20-28.
 RA Light A., du Vigneaud V.;
 RT "On the nature of oxytocin and vasopressin from human pituitary."
 RL Proc. Soc. Exp. Biol. Med. 98:692-696(1958).
 CC -1- FUNCTION: NEUROPHYSIN 1 SPECIFICALLY BINDS OXYTOCIN.
 CC -1- FUNCTION: OXYTOCIN CAUSES CONTRACTION OF THE SMOOTH MUSCLE OF THE
 CC UTERUS AND OF THE MAMMARY GLAND.
 CC -1- PHARMACEUTICAL: OXYTOCIN IS AVAILABLE UNDER THE NAMES PITOCIN
 CC (PARKE-DAVIS) AND SYNTOCINON (SANDOZ). USED TO ARTIFICIALLY SPEED
 CC OR INDUCE LABOR.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M25650; AAA59977.1; -;
DR EMBL; M11186; AAA98806.1; -;
DR EMBL; X03173; CAA26936.1; -;
DR EMBL; M62611; AAA59979.1; -;
DR PIR; B23380; NFH01.
DR PDB; 1XY1; 15-OCT-90.
DR PDB; 1XY2; 15-OCT-90.
DR MIM; 167050; -;
DR INTERPRO; IPR000981; -;
DR PFAM; PF00220; hormone4; 1.
DR PFAM; PF00184; hormone5; 1.
DR PRINTS; PR00831; NEUROPHYSIN.
DR PROSITE; PS00264; NEUROHYPOPHYS_HORM; 1.
KW Hormone; Hypothalamus; Amidation; Cleavage on pair of basic residues;
KW Signal; Pharmaceutical; 3D-structure.
FT SIGNAL 1 19
FT PEPTIDE 20 28 OXYTOCIN.
FT PEPTIDE 32 125 NEUROPHYSIN 1.
FT DISULFID 20 25
FT MOD_RES 28 28 AMIDATION (G-29 PROVIDE AMIDE GROUP).
FT DISULFID 41 85 BY SIMILARITY.
FT DISULFID 44 58 BY SIMILARITY.
FT DISULFID 59 65 BY SIMILARITY.
FT DISULFID 52 75 BY SIMILARITY.
FT DISULFID 92 104 BY SIMILARITY.
FT DISULFID 98 116 BY SIMILARITY.
FT DISULFID 105 110 BY SIMILARITY.
FT CONFLICT 94 94 S -> G (IN REF. 3).
FT CONFLICT 100 100 MISSING (IN REF. 2).
FT CONFLICT 100 101 VL -> AA (IN REF. 6).
FT CONFLICT 124 124 Q -> L (IN REF. 6).
SQ SEQUENCE 125 AA; 12722 MW; C65BB544731A0E7C CRC64;

Query Match 3.5%; Score 6; DB 1; Length 125;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 164 AGPSLA 169
Db 2 AGPSLA 7
|||||

RESULT 26
NEUL_PIG STANDARD; PRT; 125 AA.
AC P01177;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE OXYTOCIN-NEUROPHYSIN 1 PRECURSOR (OT-NPI) [CONTAINS: OXYTOCIN
DE (OXYTOCIN); NEUROPHYSIN 1].
GN OXT.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90262635; PubMed=1971513;
RA Rehbein M., Richter D.;
RT "Poly(A) tail length of oxytocin- and lysine vasopressin-encoding
RT mRNAs increases during development in the porcine hypothalamus.";
RL J. Mol. Endocrinol. 4:151-158(1990).
RN [2]
RP SEQUENCE OF 20-28.

RA Pierce J.G., Gordon S., du Vigneaud V.;
RT "Further distribution studies on the oxytocic hormone of the posterior
RT lobe of the pituitary gland and the preparation of an active
RT crystalline flavanate.";
RL J. Biol. Chem. 199:929-940(1952).
RN [3]
RP SEQUENCE OF 32-124.
RX MEDLINE=79148658; PubMed=428540;
RA Chauvet M.-T., Codogno P., Chauvet J., Acher R.;
RT "Comparison between MSEU- and VLDV-neurophysins. Complete amino acid
RT sequences of porcine and bovine VLDV-neurophysins.";
RL FEBS Lett. 98:37-40(1979).
CC -!- FUNCTION: NEUROPHYSIN 1 SPECIFICALLY BINDS OXYTOCIN.
CC -!- FUNCTION: OXYTOCIN CAUSES CONTRACTION OF THE SMOOTH MUSCLE OF THE
CC UTERUS AND OF THE MAMMARY GLAND.
DR PIR; A01439; NFPG1.
DR PIR; B60970; B60970.
DR PDB; 1XY1; 15-OCT-90.
DR PDB; 1XY2; 15-OCT-90.
DR INTERPRO; IPR000981; -;
DR PFAM; PF00220; hormone4; 1.
DR PFAM; PF00184; hormone5; 1.
DR PRINTS; PR00831; NEUROPHYSIN.
DR PROSITE; PS00264; NEUROHYPOPHYS_HORM; 1.
KW Hormone; Hypothalamus; Amidation; Cleavage on pair of basic residues;
KW Signal; 3D-structure.
FT SIGNAL 1 19
FT PEPTIDE 20 28 OXYTOCIN.
FT PEPTIDE 32 125 NEUROPHYSIN 1.
FT DISULFID 20 25 AMIDATION (G-29 PROVIDE AMIDE GROUP).
FT MOD_RES 28 28 BY SIMILARITY.
FT DISULFID 41 85 BY SIMILARITY.
FT DISULFID 44 58 BY SIMILARITY.
FT DISULFID 59 65 BY SIMILARITY.
FT DISULFID 52 75 BY SIMILARITY.
FT DISULFID 92 104 BY SIMILARITY.
FT DISULFID 98 116 BY SIMILARITY.
FT DISULFID 105 110 BY SIMILARITY.
SQ SEQUENCE 125 AA; 12887 MW; AA20808066E25B5F CRC64;
Query Match 3.5%; Score 6; DB 1; Length 125;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 164 AGPSLA 169
Db 2 AGPSLA 7
|||||

RESULT 27
NB8M_HUMAN STANDARD; PRT; 134 AA.
ID NB8M_HUMAN
AC P17568;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE NADH-UBIQUINONE OXIDOREDUCTASE B18 SUBUNIT (EC 1.6.5.3) (EC 1.6.99.3)
DE (COMPLEX I-B18) (CI-B18) (CELL ADHESION PROTEIN SQM1).
GN NDUBF7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90147818; PubMed=2302251;
RA Wong Y.-C., Tsao S.-W., Kakefuda M., Bernal S.D.;
RT "cDNA cloning of a novel cell adhesion protein expressed in human
RT squamous carcinoma cells.";
RL Biochem. Biophys. Res. Commun. 166:984-992(1990).
RN [2]
RP IDENTIFICATION OF PROBABLE FRAMESHIFTS.
RA Bairoch A.;

Unpublished observations (SEP-1997).
-!- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED TO BE UBIQUINONE.
-!- CATALYTIC ACTIVITY: NADH + UBIQUINONE -> NAD(+) + UBIQUINOL.
-!- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 40 DIFFERENT SUBUNITS.
-!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE; MATRIX SIDE.
-!- CAUTION: THE RECONSTRUCTION OF THE SEQUENCE WAS HELPED BY TWO PARTIAL EST SEQUENCES (AA548676 AND N31341). THERE ARE PROBABLY SOME MORE ERRORS IN THE C-TERMINAL PART.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M33374; AAA35675.1; -
PIR; A34653; A34653.
DR MIM; 603842; -
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Myristate.
FT INIT_MET 0 0 BY SIMILARITY.
FT LIPID 1 1 MYRISTATE (BY SIMILARITY).
FT CONFLICT 74 81 NFLACQOE -> SCWPASRK (IN REF. 1).
FT CONFLICT 85 92 WDYCEHRD -> SGLLRITAS (IN REF. 1).
SQ SEQUENCE 134 AA; 15648 MW; 4D15DAEB84AC92CF CRC64;

Query Match 3.5%; Score 6; DB 1; Length 134;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 152 RSGGRR 157
DB 110 RSGGRR 115

RESULT 28
MPI2_LYMST STANDARD; PRT; 137 AA.
AC P25289;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE MOLLUSCAN INSULIN-RELATED PEPTIDE 2 PRECURSOR (MIP II).
OS Lymnaea stagnalis (Great pond snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
OC Lymnaeidae; Lymnaea.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-CNS;
RX MEDLINE-91330924; PubMed-1868853;
RA Smit A.B., Geraerts W.P.M., Meester I., van Heerikhuizen H., Joosse J.;
RT "Characterization of a cDNA clone encoding molluscan insulin-related peptide II of Lymnaea stagnalis.";
RL Eur. J. Biochem. 199;699-703(1991).
RN [2]
RP SEQUENCE OF 32-66 AND 113-137.
RC TISSUE-LIGHT-GREEN CELLS;
RX MEDLINE-92283200; PubMed-1350761;
RA Li K.-W., Geraerts W.P.M., Joosse J.;
RT "Purification and sequencing of molluscan insulin-related peptide II from the neuroendocrine light green cells in Lymnaea stagnalis.";
RL Endocrinology 130:3427-3432(1992).
CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO DISULFIDE BONDS.
CC -!- SUBCELLULAR LOCATION: SECRETORY GRANULES.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE CEREBRAL LIGHT-GREEN CELLS WHICH ARE GIANT NEUROENDOCRINE CELLS INVOLVED IN THE CONTROL OF GROWTH.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

-!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X59302; CAA41989.1; -
PIR; S17195; S17195.
DR INTERPRO; IPR000739; -
DR PFAM; PF00049; Insulin; 1.
DR PROSITE; PS00262; INSULIN; 1.
KW Signal: Cleavage on pair of basic residues; Insulin family;
KW Multigene family.
FT SIGNAL 1 31
FT CHAIN 32 68

MOLLUSCAN INSULIN-RELATED PEPTIDE 2, B CHAIN.
FT PROPEP 71 83
FT PROPEP 86 110
FT CHAIN 113 137

MOD_RES 32 32
MOD_RES 113 113
DISULFID 47 123
DISULFID 59 136
DISULFID 122 127
BY SIMILARITY.
SQ SEQUENCE 137 AA; 15257 MW; B9E56409D1EECB41 CRC64;

Query Match 3.5%; Score 6; DB 1; Length 137;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 39 RDAETG 44
DB 70 RDAETG 75

RESULT 29
MPI5_LYMST STANDARD; PRT; 138 AA.
AC P31241;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE MOLLUSCAN INSULIN-RELATED PEPTIDE 5 PRECURSOR (MIP V).
OS Lymnaea stagnalis (Great pond snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
OC Lymnaeidae; Lymnaea.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-CNS;
RX MEDLINE-92356799; PubMed-1323019;
RA Smit A.B., Thijssen S.F.T., Geraerts W.P.M., Meester I., Heerikhuizen H., Joosse J.;
RT "Characterization of a cDNA clone encoding molluscan insulin-related peptide V of Lymnaea stagnalis.";
RL Brain Res. Mol. Brain Res. 14:7-12(1992).
CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO DISULFIDE BONDS.
CC -!- SUBCELLULAR LOCATION: SECRETORY GRANULES.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE CEREBRAL LIGHT-GREEN CELLS WHICH ARE GIANT NEUROENDOCRINE CELLS INVOLVED IN THE CONTROL OF GROWTH.

-!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

-----NOT_ANNOTATED_CDS.

DR PIR; A43957; A43957.
DR INTERPRO; IPR000739; -.
DR PFAM; PF00049; Insulin; 1.
DR PROSITE; PS00262; INSULIN; 1.
KW Signal; Cleavage on pair of basic residues; Insulin family;
KW Multigene family.
FT SIGNAL 1 31
FT CHAIN 32 69
FT FT
FT FT
FT PROPEP 72 84
FT PROPEP 87 111
FT CHAIN 114 138
FT FT
FT MOD_RES 32 32
FT FT
FT MOD_RES 114 114
FT FT
FT FT
FT DISULFID 48 124
FT DISULFID 60 137
FT DISULFID 123 128
FT DISULFID 123 128
SQ SEQUENCE 138 AA; 15540 MW; 9EF1921B84A9D6CA CRC64;

Query Match 3.5%; Score 6; DB 1; Length 138;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 RDAETG 44
DB 71 RDAETG 76

RESULT 30
PER_DROPI
ID PER_DROPI STANDARD; PRT; 143 AA.
AC Q25206;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PERIOD CIRCADIAN PROTEIN (FRAGMENT).
GN PER.
OS Drosophila picticornis (Fruit fly) (*Idiomya picticornis*).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-95115533; PubMed-7815924;
RA Nielsen J., Peixoto A.A., Piccin A., Costa R., Kyriacou C.P.,
RA Chalmers D.;
RT "Big flies, small repeats: the 'Thr-Gly' region of the period gene in
RT Diptera.";
RL Mol. Biol. Evol. 11:839-853(1994).
CC -!- FUNCTION: ESSENTIAL FOR BIOLOGICAL CLOCK FUNCTIONS. DETERMINES THE
CC PERIOD LENGTH OF CIRCADIAN AND ULTRADIAN RHYTHMS; AN INCREASE IN
CC PER DOSAGE LEADS TO SHORTENED CIRCADIAN RHYTHMS AND A DECREASE
CC LEADS TO LENGTHENED CIRCADIAN RHYTHMS. ESSENTIAL FOR THE CIRCADIAN
CC RHYTHMICITY OF LOCOMOTOR ACTIVITY, ECLOSION BEHAVIOR, AND FOR THE
CC RHYTHMIC COMPONENT OF THE MALE COURTSHIP SONG THAT ORIGINATES IN
CC THE THORACIC NERVOUS SYSTEM. THE BIOLOGICAL CYCLE DEPENDS ON THE
CC RHYTHMIC FORMATION AND NUCLEAR LOCALIZATION OF THE TIM-PER
CC COMPLEX. LIGHT INDUCES THE DEGRADATION OF TIM, WHICH PROMOTES
CC ELIMINATION OF PER. NUCLEAR ACTIVITY OF THE HETERODIMER
CC COORDINATIVELY REGULATES PER AND TIM TRANSCRIPTION THROUGH A
CC NEGATIVE FEEDBACK LOOP. BEHAVES AS A NEGATIVE ELEMENT IN CIRCADIAN
CC TRANSCRIPTIONAL LOOP. DOES NOT APPEAR TO BIND DNA, SUGGESTING
CC INDIRECT TRANSCRIPTIONAL INHIBITION (BY SIMILARITY).
CC -!- SUBUNIT: FORMS HETERODIMER WITH TIMELESS (TIM); THE COMPLEX THEN

CC TRANSLOCATES INTO THE NUCLEUS (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR AT SPECIFIC PERIODS OF THE DAY.
CC FIRST ACCUMULATES IN THE PERINUCLEAR REGION ABOUT ONE HOUR BEFORE
CC TRANSLOCATION INTO THE NUCLEUS. INTERACTION WITH TIM IS REQUIRED
CC FOR NUCLEAR LOCALIZATION (BY SIMILARITY).
CC -!- PTM: PHOSPHORYLATED WITH A CIRCADIAN RHYTHMICITY, PROBABLY BY THE
CC DOUBLE-TIME PROTEIN (DBT). PHOSPHORYLATION COULD BE IMPLICATED IN
CC THE STABILITY OF PER MONOMER AND IN THE FORMATION OF HETERODIMER
CC PER-TIM (BY SIMILARITY).
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U11810; AAA76592.1; -.
DR FLYBASE; FBgn0015136; Dpic\per.
KW Biological rhythms; Repeat; Nuclear protein; Phosphorylation.
FT NON_TER 1
FT NON_TER 143 143
SQ SEQUENCE 143 AA; 14459 MW; 5F4F2DC27F5D6846 CRC64;

Query Match 3.5%; Score 6; DB 1; Length 143;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 AGPSLA 169
DB 120 AGPSLA 125

RESULT 31
CNRD_BOVIN
ID CNRD_BOVIN STANDARD; PRT; 150 AA.
AC Q95142;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE RETINAL ROD RHODOPSIN-SENSITIVE CGMP 3',5'-CYCLIC PHOSPHODIESTERASE
DE DELTA-SUBUNIT (EC 3.1.4.17) (GMP-PDE DELTA).
GN PDE6D.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE-RETINA;
RX MEDLINE-96394534; PubMed-8798640;
RA Florio S.K., Prusti R.K., Beavo J.A.;
RT "Solubilization of membrane-bound rod phosphodiesterase by the rod
RT phosphodiesterase recombinant delta subunit.";
RL J. Biol. Chem. 271:24036-24047(1996).
CC -!- CATALYTIC ACTIVITY: GUANOSINE 3',5'-CYCLIC PHOSPHATE + H(2)O =
CC GUANOSINE 5'-PHOSPHATE.
CC -!- SUBUNIT: OLIGOMER COMPOSED OF TWO CATALYTIC CHAINS (ALPHA AND
CC BETA), AN INHIBITORY CHAIN (GAMMA) AND THE DELTA CHAIN.
CC -!- TISSUE SPECIFICITY: RETINA.
CC -!- SIMILARITY: BELONGS TO THE PDE6D / UNC119 FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; U65073; AAC48634.1; -.

```
KW Hydrolase; cGMP; Vision.
SQ SEQUENCE 150 AA; 17390 MW; 5E8D881372905EA0 CRC64;

Query Match 3.5%; Score 6; DB 1; Length 150;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 RDAETG 44
DB 23 RDAETG 28

RESULT 32
CNRD_CANFA STANDARD; PRT; 150 AA.
AC G9XT54;
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE RETINAL ROD RHODOPSIN-SENSITIVE CGMP 3',5'-CYCLIC PHOSPHODIESTERASE
DE DELTA-SUBUNIT (EC 3.1.4.17) (GMP-PDE DELTA).
GN PDE6D.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=99384013; PubMed=10452952;
RA Wang W., Zhang Q., Acland G.M., Mellersh C., Ostrander E.A., Ray K.,
RA Aquirre G.D.;
RT "Molecular characterization and mapping of canine cGMP-
RT phosphodiesterase delta subunit (PDE6D).";
RL Gene 236:325-332(1999).
CC -1- CATALYTIC ACTIVITY: GUANOSINE 3',5'-CYCLIC PHOSPHATE + H(2)O =
CC GUANOSINE 5'-PHOSPHATE.
CC -1- SUBUNIT: OLIGOMER COMPOSED OF TWO CATALYTIC CHAINS (ALPHA AND
CC BETA), AN INHIBITORY CHAIN (GAMMA) AND THE DELTA CHAIN.
CC -1- TISSUE SPECIFICITY: RETINA.
CC -1- SIMILARITY: BELONGS TO THE PDE6D / UNC119 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF113996; AAD39365.1; -
CC EMBL; AF109151; AAD54058.1; -
CC Hydrolase; cGMP; Vision.
SQ SEQUENCE 150 AA; 17362 MW; 797002FC4F505EBF CRC64;

Query Match 3.5%; Score 6; DB 1; Length 150;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 RDAETG 44
DB 23 RDAETG 28

RESULT 33
CNRD_HUMAN STANDARD; PRT; 150 AA.
AC O43924; O43250;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE RETINAL ROD RHODOPSIN-SENSITIVE CGMP 3',5'-CYCLIC PHOSPHODIESTERASE
DE DELTA-SUBUNIT (EC 3.1.4.17) (GMP-PDE DELTA) (P17 PROTEIN).
GN PDE6D OR PDED.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=98234546; PubMed=9570951;
RA Li N., Florio S.K., Pettenati M.J., Rao P.N., Beavo J.A., Baehr W.;
RT "Characterization of human and mouse rod cGMP phosphodiesterase delta
RT subunit (PDE6D) and chromosomal localization of the human gene.";
RL Genomics 49:76-82(1998).
[2]
RN SEQUENCE FROM N.A.
RP MEDLINE=98194225; PubMed=9533031;
RA Erchoya G., Derre J., Chatelin S., Nancy V., Berger R., Kaplan J.,
RA Munnich A., de Gunzburg J.;
RT "cDNA sequence, genomic organization and mapping of PDE6D, the human
RT gene encoding the delta subunit of the cGMP phosphodiesterase of
RT retinal rod cells to chromosome 2q36.";
RL Cytogenet. Cell Genet. 79:139-141(1997).
[3]
RN SEQUENCE FROM N.A.
RP MEDLINE=98454318; PubMed=9781033;
RA Lorenz B., Migliaccio C., Lichtner P., Meyer C., Strom T.M.,
RA D'Urso M., Becker J., Ciccodicola A., Meitinger T.;
RT "Cloning and gene structure of the rod cGMP phosphodiesterase delta
RT subunit gene (PDED) in man and mouse.";
RL Eur. J. Hum. Genet. 6:283-290(1998).
[4]
RN SEQUENCE FROM N.A.
RP MEDLINE=98454318; PubMed=9781033;
RA Zahraoui A., Marzocco A.M., Galli T., Louvard D.;
RT "A novel RAB interacting protein.";
RA Submitted (SEP-1997) to the EMBL/GenBank/DBDJ databases.
CC -1- CATALYTIC ACTIVITY: GUANOSINE 3',5'-CYCLIC PHOSPHATE + H(2)O =
CC GUANOSINE 5'-PHOSPHATE.
CC -1- SUBUNIT: OLIGOMER COMPOSED OF TWO CATALYTIC CHAINS (ALPHA AND
CC BETA), AN INHIBITORY CHAIN (GAMMA) AND THE DELTA CHAIN.
CC -1- TISSUE SPECIFICITY: RETINA.
CC -1- SIMILARITY: BELONGS TO THE PDE6D / UNC119 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF045999; AAC39720.1; -
CC EMBL; AF022912; AAB87872.1; -
CC EMBL; AF042835; AAC25953.1; -
CC EMBL; AF042833; AAC25953.1; JOINED.
CC EMBL; AF042834; AAC25953.1; JOINED.
CC EMBL; AJ001626; CAA04880.1; -
CC MIM; 602676; -
CC Hydrolase; cGMP; Vision.
FT CONFLICT 117 117 M -> V (IN REF. 4).
FT CONFLICT 146 146 R -> G (IN REF. 4).
SQ SEQUENCE 150 AA; 17420 MW; AB8D9309C33B4411 CRC64;

Query Match 3.5%; Score 6; DB 1; Length 150;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 RDAETG 44
DB 23 RDAETG 28

RESULT 34
CNRD_MOUSE STANDARD; PRT; 150 AA.
ID CNRD_MOUSE
```


055057;
15-JUL-1999 (Rel. 38, Created)
15-JUL-1999 (Rel. 38, Last sequence update)
01-OCT-2000 (Rel. 40, Last annotation update)
RETINAL ROD RHODOPSIN-SENSITIVE CGMP 3',5'-CYCLIC PHOSPHODIESTERASE
DELTA-SUBUNIT (EC 3.1.4.17) (GMP-PDE DELTA).
PDE6D.
Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
[1]
SEQUENCE FROM N.A.
RP MEDLINE=98234546; PubMed=9570951;
Li N., Florio S.K., Pattenati M.J., Rao P.N., Beavo J.A., Baehr W.;
"Characterization of human and mouse rod cGMP phosphodiesterase delta
subunit (PDE6D) in man and mouse.";
Eur. J. Hum. Genet. 6:283-290(1998).
CC -!- CATALYTIC ACTIVITY: GUANOSINE 3',5'-CYCLIC PHOSPHATE + H(2)O =
GUANOSINE 5'-PHOSPHATE.
CC -!- SUBUNIT: OLIGOMER COMPOSED OF TWO CATALYTIC CHAINS (ALPHA AND
BETA), AN INHIBITORY CHAIN (GAMMA) AND THE DELTA CHAIN.
CC -!- TISSUE SPECIFICITY: RETINA.
CC -!- SIMILARITY: BELONGS TO THE PDE6D / UNC119 FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AF046000; AAC40094.1; -;
DR EMBL; AF039216; AAC25686.1; -;
DR MGD; MGI:1270843; PDE6D.
KW Hydrolase; cGMP; Vision.
SQ SEQUENCE 150 AA; 17348 MW; 5E8DB9C8020E38D5 CRC64;

Query Match 3.5%; Score 6; DB 1; Length 150;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 RDAETG 44
Db 23 RDAETG 28
|||||

RESULT 35
YK91_CAEEL STANDARD; PRT; 159 AA.
AC Q18268;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HYPOTHEICAL 18.4 KDA PROTEIN C27H5.1 IN CHROMOSOME II.
GN C27H5.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Pauley A.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: BELONGS TO THE PDE6D / UNC119 FAMILY. STRONG, TO
MAMMALIAN RETINAL ROD RHODOPSIN-SENSITIVE CGMP 3',5'-CYCLIC
PHOSPHODIESTERASE DELTA-SUBUNIT.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; U14635; AAC46655.1; -;
DR WORMPEP; C27H5.1; C501175.
KW Hypothetical protein.
SQ SEQUENCE 159 AA; 18372 MW; EBB868F29F47FD59 CRC64;

Query Match 3.5%; Score 6; DB 1; Length 159;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 RDAETG 44
Db 32 RDAETG 37
|||||

RESULT 36
UL20_PRVN3 STANDARD; PRT; 161 AA.
ID UL20_PRVN3
AC Q00702;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE UL20 MEMBRANE PROTEIN HOMOLOG.
OS Pseudorabies virus (strain NIA-3) (PRV).
CC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
CC Alphaherpesvirinae; Varicellovirus.
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93059656; PubMed=1331512;
RA Berns A.A., Kimman T.T., Pol J.J., Wagenaar F.F., de Wind N.N.;
"The pseudorabies virus homology of the herpes simplex virus UL21
gene product is a capsid protein which is involved in capsid
maturation."
RL J. Virol. 66:7096-7103(1992).
CC -!- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED OR TRANSMEMBRANE
PROTEIN (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL20,
EHV-1 41, AND VZV 39.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; M95285; AAA47473.1; -;
DR Transmembrane.
SQ SEQUENCE 161 AA; 16640 MW; 218083FDEB74AA45 CRC64;

Query Match 3.5%; Score 6; DB 1; Length 161;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 CLVLAL 18
Db 57 CLVLAL 62
|||||

```
RESULT 37
VPB_BPP2
ID VPB_BPP2 STANDARD; PRT; 166 AA.
AC P07696;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE REPLICATION GENE B PROTEIN (GPB).
GN B.
OS Bacteriophage P2.
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Myoviridae.
RN [1]
RP MEDLINE=87286423; PubMed=3475531;
RX Haggard-Ljungquist E., Kockum K., Bertani L.E.;
RT "DNA sequences of bacteriophage P2 early genes cox and B and their
RT regulatory sites."
RL Mol. Gen. Genet. 208:52-56(1987).
CC -!- FUNCTION: REQUIRED FOR LAGGING STRAND SYNTHESIS. MIGHT INTERACT
CC WITH THE HOST DNAB PROTEIN.
CC -!- DEVELOPMENTAL STAGE: REQUIRED CONTINUOUSLY THROUGHOUT THE
CC INFECTION CYCLE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF063097; AAD03301.1;
DR PIR; S07044; S07044.
KW Early protein; DNA replication; DNA synthesis.
SQ SEQUENCE 166 AA; 19887 MW; 8C9C64338AAB9C20 CRC64;

Query Match 3.5%; Score 6; DB 1; Length 166;
Best Local Similarity 100.0%; Pred. No. le+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

QY 45 ERLVCA 50
Db 71 ERLVCA 76

RESULT 38
YHIL_AZOBR
ID YHIL_AZOBR STANDARD; PRT; 168 AA.
AC P26723;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE HYPOTHETICAL 16.6 KDA PROTEIN IN HISH-HISA INTERGENIC REGION (ORF1).
OS Azospirillum brasilense.
OC Bacteria; Proteobacteria; alpha subdivision; Rhodospirillaceae;
OC Azospirillum.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP6;
RX MEDLINE=89313660; PubMed=2664449;
RA Fani R., Allotta G., Bazzicalupo M., Gallori E., Polsinelli M.;
RT "Cloning of histidine genes of Azospirillum brasilense: organization
RT of the ABFH gene cluster and nucleotide sequence of the hisB gene.";
RL Mol. Gen. Genet. 216:224-229(1989).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```
CC EMBL; X61207; CAA43517.1;
DR PIR; S16800; S16800.
DR INTERPRO; IPR000182;
DR PFAM; PF00583; Acetyltransf; 1.
KW Hypothetical protein; Histidine biosynthesis.
SQ SEQUENCE 168 AA; 18550 MW; ACE1340C18A21116 CRC64;

Query Match 3.5%; Score 6; DB 1; Length 168;
Best Local Similarity 100.0%; Pred. No. le+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

QY 99 EEEARA 104
Db 114 EEEARA 119

RESULT 39
YC50_AQUAE
ID YC50_AQUAE STANDARD; PRT; 171 AA.
AC O67293;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYPOTHETICAL PROTEIN AQ_1250.
GN AQ_1250.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VFS;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., AuJay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus."
RL Nature 392:353-358(1998).
CC -!- SIMILARITY: BELONGS TO THE UPF0098 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000730; AAC07255.1;
DR Hypothetical protein.
SQ SEQUENCE 171 AA; 18917 MW; A1F738752E15FF71 CRC64;

Query Match 3.5%; Score 6; DB 1; Length 171;
Best Local Similarity 100.0%; Pred. No. le+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

QY 71 GPCPPR 76
Db 115 GPCPPR 120

RESULT 40
YHL4_HCMVA
ID YHL4_HCMVA STANDARD; PRT; 172 AA.
AC P09698;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-AUG-1990 (Rel. 15, Last annotation update)
DE HYPOTHETICAL PROTEIN HHLF4.
OS Human cytomegalovirus (strain AD169).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
```

```

OC Betaherpesvirinae; Cytomegalovirus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87169717; PubMed=3031311;
RA Weston K., Barrell B.G.;
RT "Sequence of the short unique region, short repeats, and part of the
RT long repeats of human cytomegalovirus.";
RL J. Mol. Biol. 192:177-208(1986).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X04650; CAB37118.1; -.
DR PIR; G27216; QQBED7.
KW Hypothetical protein.
SQ SEQUENCE 172 AA; 18223 MW; 0E50EB90E011CA9F CRC64;

Query Match 3.5%; Score 6; DB 1; Length 172;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 29 GVAETP 34
Db 77 GVAETP 82

RESULT 41
YHNN_CAEEL STANDARD; PRT; 189 AA.
AC Q94402;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE HYPOTHETICAL 22.3 KDA PROTEIN ZK265.6 IN CHROMOSOME I.
GN ZK265.6.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Caenorhabditis.
OC Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RA Dobson R.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: STRONG, TO HUMAN CGI-117.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z81143; CAB03516.1; -.
DR WORMPEP; ZK265.6; CE15299.
KW Hypothetical protein.
SQ SEQUENCE 189 AA; 22259 MW; F7619CF2A7FA1F34 CRC64;

Query Match 3.5%; Score 6; DB 1; Length 189;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 99 EEEARA 104
Db 120 EEEARA 125

```

```

RESULT 42
CLD7_RAT
ID CLD7_RAT STANDARD; PRT; 191 AA.
AC Q9ZILL;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE CLAUDIN-7 (FRAGMENT).
GN CLDN7.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RA Keen T.J., Inglehearn C.F.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: COMPONENT OF TIGHT JUNCTION (TJ) STRANDS.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE CLAUDIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ011811; CAA09790.1; -.
DR INTERPRO; IPR000729; -.
DR INTERPRO; IPR001832; -.
DR PFAM; PF008822; PMP22_Claudin; 1.
DR PRINTS; PR01077; CLAUDIN.
DR PROSITE; PS01346; CLAUDIN; 1.
KW Tight junction; Transmembrane.
FT NON_TER 1
FT TRANSMEM 62 82 POTENTIAL.
FT TRANSMEM 100 120 POTENTIAL.
FT TRANSMEM 141 161 POTENTIAL.
SQ SEQUENCE 191 AA; 20366 MW; 83B445908DFFF41A CRC64;

Query Match 3.5%; Score 6; DB 1; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLALPA 20
Db 50 VLALPA 55

RESULT 43
YGL2_STRCO
ID YGL2_STRCO STANDARD; PRT; 191 AA.
AC P40182;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DE HYPOTHETICAL 20.1 KDA PROTEIN IN GLK 5'REGION (ORF2).
GN SC6E10.21C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=A3(2);
RX MEDLINE=93062017; PubMed=1435260;
RA Angell S., Schwarz E., Bibb M.J.;
RT "The glucose kinase gene of Streptomyces coelicolor A3(2): its
RT nucleotide sequence, transcriptional analysis and role in glucose
RT repression.";
RL Mol. Microbiol. 6:2833-2844(1992).

```

```

[2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Seeger K., Harris D., James K.D., Parkhill J., Barrell B.G.,
RA Rajadream M.A.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INVOLVED IN GLUCOSE TRANSPORT OR METABOLISM.
CC -!- SUBUNIT: IT MAY FORM AN HETEROTETRAMER OF TWO GLUCOKINASE SUBUNITS
CC WITH TWO ORF2 PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X65932; CAA46726.1; -.
DR EMBL: AL109661; CAB51975.1; -.
DR PIR: S26207; S26207.
KW Hypothetical protein; Glucose metabolism; Sugar transport.
SQ SEQUENCE 191 AA; 20131 MW; 7CF36B2A7AC20044 CRC64;
-----
Query Match 3.5%; Score 6; DB 1; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 161 GQVAGP 166
Db 93 GQVAGP 98
|||||
-----
RESULT 44
YC69_PYRHO
ID YC69_PYRHO STANDARD; PRT; 198 AA.
AC O58984;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYPOTHETICAL PROTEIN PH1269.
GN PH1269.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).
CC -!- SIMILARITY: BELONGS TO THE UPF0098 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF000005; BAA30372.1; -.
DR Hypothetical protein.
SQ SEQUENCE 198 AA; 21690 MW; 83C1B9E29E2BCF03 CRC64;
-----

```

```

Query Match 3.5%; Score 6; DB 1; Length 198;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 71 GPCPPR 76
Db 143 GPCPPR 148
|||||
-----
RESULT 45
CTFL_RAT
ID CTFL_RAT STANDARD; PRT; 203 AA.
AC Q63086;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE CARDIOTROPIN-1 (CT-1).
DE CTFL.
GN CTFL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR; TISSUE=HEART;
RX MEDLINE=96193659; PubMed=8604995;
RA Ishikawa M., Saito Y., Miyamoto Y., Kuwahara K., Ogawa E.,
RA Nakagawa O., Harada M., Masuda I., Nakao K.;
RT "cDNA cloning of rat cardiostrophin-1 (CT-1): augmented expression of
RT CT-1 gene in ventricle of genetically hypertensive rats.";
RL Biochem. Biophys. Res. Commun. 219:377-381(1996).
CC -!- FUNCTION: INDUCES CARDIAC MYOCYTE HYPERTROPHY IN VITRO. BINDS TO
CC AND ACTIVATES THE LEUKEMIA INHIBITORY FACTOR RECEPTOR (LIF
CC RECEPTOR)/GP 130 RECEPTOR COMPLEX.
CC -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE VENTRICLE AND ATRIUM OF ADULT
CC RATS. ALSO DETECTED IN THE LUNG, KIDNEY, LIVER, SKELETAL MUSCLE,
CC STOMACH AND URINARY BLADDER. NOT DETECTED IN BRAIN, COLON, TESTIS,
CC SPLEEN OR THYMUS. OVEREXPRESSED IN THE VENTRICLES IN THE CASE OF
CC HYPERTENSION AND HYPERTROPHY.
CC -!- SIMILARITY: BELONGS TO THE IL-6 SUPERFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D78591; BAAL1427.1; -.
DR Cytokine.
SQ SEQUENCE 203 AA; 21439 MW; DF8B921A2FA1C832 CRC64;
-----
Query Match 3.5%; Score 6; DB 1; Length 203;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 ALPALL 22
Db 96 ALPALL 101
|||||
-----
RESULT 46
IDI_MYCTU
ID IDI_MYCTU STANDARD; PRT; 203 AA.
AC P72002; O08150;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PROBABLE ISOPENTENYL-DIPHOSPHATE DELTA-ISOMERASE (EC 5.3.3.2) (IPP
DE ISOMERASE) (ISOPENTENYL PYROPHOSPHATE ISOMERASE).
GN IDI OR RV1745C OR MTCY28.08C OR MTCY04C12.29C.
-----

```

```

OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies K., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
RA Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
CC -1- FUNCTION: CATALYZES THE 1,3-ALLYLIC REARRANGEMENT OF THE
CC HOMOALLYLIC SUBSTRATE ISOPENTENYL (IPP) TO ITS ALLYLIC ISOMER,
CC DIMETHYLLALLYL DIPHOSPHATE (DMAPP) (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ISOPENTENYL DIPHOSPHATE - DIMETHYLLALLYL
CC DIPHOSPHATE.
CC -1- SIMILARITY: BELONGS TO THE IPP ISOMERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z95890; CAB09331.1; -.
CC TUBERCULIST; RV1745C; -.
CC INTERPRO; IPR002667; -.
CC PFAM; PF01772; IPP isomerase; 1.
CC Isomerase; Isoprene biosynthesis.
CC ACT_SITE 75 75 BY SIMILARITY.
CC ACT_SITE 124 124 BY SIMILARITY.
CC SEQUENCE 203 AA; 22489 MW; B5F1IA06B13B6747 CRC64;

Query Match 3.5%; Score 6; DB 1; Length 203;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 136 AGVIAP 141
Db 161 AGVIAP 166
|||||

RESULT 47
FTSQ_AGRU
ID FTSQ_AGRU STANDARD; PRT; 210 AA.
AC Q30990;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CELL DIVISION PROTEIN FTSQ HOMOLOG (FRAGMENT).
GN FTSQ.
OS Agrobacterium tumefaciens.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Agrobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=A136;
RC MEDLINE=98012980; PubMed=9352931;
RA Ma X., Sun Q., Wang R., Singh G., Jonietz E.L., Margolin W.;
RA "Interactions between heterologous FtsA and FtsZ proteins at the FtsZ
RT ring.";
RL J. Bacteriol. 179:6788-6797(1997).
CC -1- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN SEPTUM FORMATION (BY

SIMILARITY).
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. INNER MEMBRANE
CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE FTSQ FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF024659; AAC45819.1; -.
CC Cell division; Septation; Transmembrane; Inner membrane.
CC NON_TER 1
CC SEQUENCE 210 AA; 23552 MW; C1F63ED0DE6EBB21 CRC64;

Query Match 3.5%; Score 6; DB 1; Length 210;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 RDAETG 44
Db 97 RDAETG 102
|||||

RESULT 48
LAG1_HUMAN
ID LAG1_HUMAN STANDARD; PRT; 210 AA.
AC Q75638; Q75637;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE LAGE-1 PROTEIN.
GN LAGE1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
TX TISSUE=MELANOMA;
RX MEDLINE=98289662; PubMed=9626360;
RA Lethe B., Lucas S., Michaux L., de Smet C., Godelaine D., Serrano A.,
RA de Plaen E., Boon T.;
RT "LAGE-1, a new gene with tumor specificity.";
RL Int. J. Cancer 76:903-908(1998).
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; LAGE-1A AND LAGE-1B (SHOWN
CC HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: TESTIS AND VERY LOW LEVEL IN PLACENTA AND IN
CC SOME UTERUS SAMPLES. OBSERVED IN 25-50% OF TUMOR SAMPLES OF
CC MELANOMAS, NON-SMALL-CELL LUNG CARCINOMAS, BLADDER, PROSTATE AND
CC HEAD AND NECK CANCERS.
CC -1- DOMAIN: A TRANSMEMBRANE DOMAIN IS PRESENT IN ISOFORM LAGE-1A.
CC -1- SIMILARITY: STRONG, TO NY-ESO-1/CTAG.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ223093; CAAL1117.1; -.
CC EMBL; AJ223093; CAAL1116.1; -.
CC EMBL; AJ223040; CAAL1043.1; -.
CC EMBL; AJ223041; CAAL1044.1; -.
CC Polymorphism; Alternative splicing; Transmembrane; Antigen.
CC DOMAIN 5 79 GLY-RICH.
CC FT DOMAIN 183 188 POLY-PRO.
CC FT VARSPIC 135 210 MSYWDQREGAMRVVGVGIGSASPEGKARDLTPKHV

```

```
FT SEQRPGTPGPPPPGCAQDGRGVAFNMFSAPIH -> IR
FT LTAADHROLQSLSSICLQSLMLWITQCFLPFLAQAQPSG
FT QRR (IN ISOFORM LAGE-1A).
FT R -> Q.
FT /FTId=VAR_007855.
FT E -> Q.
FT /FTId=VAR_007856.
FT W -> R.
FT /FTId=VAR_007857.
FT 8BE0EE0AE55E8BE CRC64;
SQ SEQUENCE 210 AA; 21119 MW; 21119 MW; 8BE0EE0AE55E8BE CRC64;

Query Match 3.5%; Score 6; DB 1; Length 210;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 147 RGGAPR 152
Db 57 RGGAPR 62

RESULT 49
COX2_LUMTE STANDARD; PRT; 228 AA.
AC Q37545;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE II (EC 1.9.3.1).
GN COII.
OS Lumbricus terrestris (Common earthworm).
OG Mitochondrion.
OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
OC Lumbicina; Lumbricidae; Lumbricus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96042914; PubMed=8536978;
RA Boore J.L., Brown W.M.;
RT "Complete sequence of the mitochondrial DNA of the annelid worm
  Lumbricus terrestris."
RL Genet. 141:305-319(1995).
CC -1- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY
  CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNIT 1-
  3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. SUBUNIT 2
  TRANSFERS THE ELECTRONS FROM CYTOCHROME C VIA ITS BINUCLEAR COPPER
  A CENTER TO THE BIMETALLIC CENTER OF THE CATALYTIC SUBUNIT 1.
CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
  4 FERRICYTOCHROME C.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
  INNER MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 2 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC EMBL; U24570; AAC46865.1; -.
CC HSSP; P08306; IARI.
CC INTERPRO; IPR001505; -.
CC INTERPRO; IPR002429; -.
CC PFAM; PF00116; COX2; 1.
CC PRINTS; PR01166; CYCOXIDASEII.
CC PROSITE; PS00078; COX2; 1.
KW Oxidoreductase; Copper; Mitochondrion; Transmembrane; Inner membrane;
  Electron transport; Respiratory chain.
FT DOMAIN 1 26 MITOCHONDRIAL INTERMEMBRANE (POTENTIAL).
FT TRANSMEM 27 47 POTENTIAL.
FT DOMAIN 48 60 MITOCHONDRIAL MATRIX (POTENTIAL).
FT TRANSMEM 61 81 POTENTIAL.
```

```
FT DOMAIN 82 228 MITOCHONDRIAL INTERMEMBRANE (POTENTIAL).
FT METAL 161 161 COPPER A (PROBABLE).
FT METAL 196 196 COPPER A (PROBABLE).
FT METAL 200 200 COPPER A (PROBABLE).
FT METAL 204 204 COPPER A (PROBABLE).
SQ SEQUENCE 228 AA; 25659 MW; ALCE3834F3BD80D3 CRC64;

Query Match 3.5%; Score 6; DB 1; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 LVLALP 19
Db 74 LVLALP 79

RESULT 50
GPHC_ALCEU STANDARD; PRT; 231 AA.
ID GPHC_ALCEU
AC P40852;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE PHOSPHOGLYCOLATE PHOSPHATASE, CHROMOSOMAL (EC 3.1.3.18) (PGP).
GN CBBZC.
OS Alkaligenes eutrophus (Ralstonia eutropha).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=H16 / ATCC 17699;
RX MEDLINE=94042908; PubMed=8226680;
RA Schaeferjohann J., Yoo J.-G., Kusian B., Bowien B.;
RT "The cbb operons of the facultative chemototroph Alkaligenes
  eutrophus encode phosphoglycolate phosphatase."
RL J. Bacteriol. 175:7329-7340(1993).
CC -1- CATALYTIC ACTIVITY: 2-PHOSPHOGLYCOLATE + H(2)O = GLYCOLATE +
  ORTHOPHOSPHATE.
CC -1- PATHWAY: GLYCOLATE OXIDATION PATHWAY.
CC -1- SUBUNIT: HOMOTRIMER (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE CBBY/CBBZ/GPH/YIEH FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC EMBL; M68904; AAA20197.1; -.
CC INTERPRO; IPR001454; -.
CC PFAM; PF00702; Hydrolase; 1.
KW Carbohydrate metabolism; Calvin cycle; Hydrolase.
SQ SEQUENCE 231 AA; 23770 MW; EC8A8C42060066D8 CRC64;

Query Match 3.5%; Score 6; DB 1; Length 231;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 ALPALL 22
Db 217 ALPALL 222
```

Search completed: January 30, 2001, 17:10:47
Job time: 231 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 17:10:38 ; Search time 42.13 Seconds
(without alignments)
472.949 Million cell updates/sec

Title: US-09-518-931-4

Perfect score: 170

Sequence: 1 MRALGPGLSLCLVLALPA.....PRSGRRRCRGQVAGPSLAP 170

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 374700 seqs, 117207915 residues

Word size : 6

Total number of hits satisfying chosen parameters: 872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phage.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	142	83.5	300	4	Q95407 homo sapien
2	9	5.3	561	10	Q9shb2 arabidopsis
3	8	4.7	440	2	Q9jp96 rhodocyclis
4	7	4.1	82	5	O76201 phoneutria
5	7	4.1	83	12	O69290 gallid herp
6	7	4.1	106	2	Q9peh3 xylella fas
7	7	4.1	115	6	Q29173 sus scrofa
8	7	4.1	123	3	Q9uvf0 yarrowia li
9	7	4.1	124	5	Q9VEK1
10	7	4.1	173	2	Q9RJG5
11	7	4.1	239	2	Q9L7T6
12	7	4.1	266	10	Q9SF29
13	7	4.1	271	5	O44077
14	7	4.1	274	2	Q9ZDM1
15	7	4.1	286	10	Q9LGS9
16	7	4.1	292	2	Q9X9X2
17	7	4.1	293	5	O15988
18	7	4.1	294	10	Q9SM72
19	7	4.1	299	5	Q9NNF4

20	7	4.1	314	5	O15987	O15987 patinopecte
21	7	4.1	329	2	O45987	O45987 caulobacter
22	7	4.1	335	2	Q9PJX7	Q9pjx7 chlamydia m
23	7	4.1	335	2	Q9Z808	Q9z808 chlamydia p
24	7	4.1	338	2	Q9Z954	Q9z954 chlamydia p
25	7	4.1	338	5	Q9NAL9	Q9nal9 caenorhabdi
26	7	4.1	343	2	Q9K2C3	Q9k2c3 chlamydia p
27	7	4.1	360	11	Q9JL20	Q9jl20 mus musculu
28	7	4.1	361	4	Q9NZG2	Q9nzg2 homo sapien
29	7	4.1	362	11	Q9JL21	Q9jil21 mus musculu
30	7	4.1	362	11	Q9JIP1	Q9jip1 mus musculu
31	7	4.1	364	3	Q94124	Q94124 rhizopus st
32	7	4.1	374	2	Q9RJF6	Q9rjf6 streptomyce
33	7	4.1	393	1	Q9YBB6	Q9ybb6 aeropyrum p
34	7	4.1	407	2	Q9ZCV6	Q9zcv6 rickettsia
35	7	4.1	414	10	Q9M231	Q9m231 arabidopsis
36	7	4.1	419	2	Q9KX36	Q9kx36 methylocyst
37	7	4.1	432	2	Q9X804	Q9x804 streptomyce
38	7	4.1	433	2	Q9X8V6	Q9x8v6 streptomyce
39	7	4.1	446	13	P79825	P79825 oncorhynch
40	7	4.1	453	2	O66700	O66700 aquifex aeo
41	7	4.1	472	5	Q20406	Q20406 caenorhabdi
42	7	4.1	481	2	Q9R7S6	Q9r7s6 escherichia
43	7	4.1	485	2	Q9L2L1	Q9l2l1 streptomyce
44	7	4.1	501	2	Q9R2U8	Q9r2u8 streptomyce
45	7	4.1	509	10	Q9SN03	Q9sn03 arabidopsis
46	7	4.1	513	2	O52237	O52237 myxococcus
47	7	4.1	518	10	Q9SW67	Q9sw67 arabidopsis
48	7	4.1	546	2	Q9ZC50	Q9zcs0 yersinia pe
49	7	4.1	560	5	Q9NMV7	Q9nmv7 leishmania
50	7	4.1	600	2	P74569	P74569 synechocyst
51	7	4.1	607	10	Q9LGS5	Q9lgs5 oryza sativ
52	7	4.1	623	2	O30589	O30589 streptomyce
53	7	4.1	659	3	Q12256	Q12256 saccharomyc
54	7	4.1	708	2	O46444	O46444 comamonas t
55	7	4.1	734	10	Q9MAP4	Q9map4 arabidopsis
56	7	4.1	832	2	Q9R7S7	Q9r7s7 escherichia
57	7	4.1	914	2	O53964	O53964 streptomyce
58	7	4.1	1063	5	Q9KDU5	Q9kdu5 bacillus ha
59	7	4.1	1180	5	Q9U143	Q9u143 leishmania
60	7	4.1	1268	2	Q9K2L1	Q9k2l1 streptomyce
61	7	4.1	1289	11	Q9WTR2	Q9wtr2 mus musculu
62	7	4.1	1326	2	Q9L2C3	Q9l2c3 streptomyce
63	7	4.1	1525	2	Q9L0T6	Q9l0t6 streptomyce
64	7	4.1	1611	2	O33957	O33957 streptomyce
65	7	4.1	1681	11	O62467	O62467 mus musculu
66	7	4.1	1980	5	Q9U109	Q9u109 leishmania
67	7	4.1	2187	11	P70670	P70670 mus musculu
68	7	4.1	2786	5	Q9NE64	Q9ne64 leishmania
69	7	4.1	4599	4	Q9NZR2	Q9nzz2 homo sapien
70	7	4.1	4599	11	Q9JL18	Q9jil18 mus musculu
71	6	3.5	20	4	Q9UNE8	Q9une8 homo sapien
72	6	3.5	23	2	Q9R4C9	Q9r4c9 pseudomonas
73	6	3.5	23	12	Q87081	Q87081 pseudorabie
74	6	3.5	25	11	Q9JKV4	Q9jkv4 mus musculu
75	6	3.5	40	2	O46498	O46498 desulfovibr
76	6	3.5	43	4	Q14601	Q14601 homo sapien
77	6	3.5	56	11	O61725	O61725 mus musculu
78	6	3.5	75	5	O00927	O00927 trypanosoma
79	6	3.5	80	2	Q9ZAX9	Q9zax9 phingomona
80	6	3.5	80	5	P81793	P81793 phoneutria
81	6	3.5	82	6	Q9XSS3	Q9xss3 canis famli
82	6	3.5	83	10	Q9SHW8	Q9shw8 arabidopsis
83	6	3.5	84	5	P91899	P91899 bombyx mori
84	6	3.5	88	12	Q9QPR8	Q9qpr8 hepatitis c
85	6	3.5	96	4	Q9Y3X1	Q9y3x1 homo sapien
86	6	3.5	99	12	Q9YMM6	Q9ymm6 lymantria d
87	6	3.5	100	4	Q9UGS8	Q9ugs8 homo sapien
88	6	3.5	101	2	Q9LBN1	Q9lbn1 uncultured
89	6	3.5	102	2	O86496	O86496 streptomyce
90	6	3.5	102	12	O68358	O68358 hepatitis c
91	6	3.5	103	2	O50015	O50015 mycobacteri
92	6	3.5	105	12	O56310	O56310 egyptian su

93	105	12	Q9YXE2	Q9YXE2 egyptian su	166	6	3.5	181	2	O82986	O82986 bacillus sp
94	105	12	Q9YXD8	Q9YXD8 egyptian su	167	6	3.5	186	1	Q9UZJ3	Q9UZJ3 pyrococcus
95	105	12	Q9YXD5	Q9YXD5 egyptian su	168	6	3.5	189	4	Q9Y518	Q9Y518 homo sapien
96	105	12	Q9IH25	Q9IH25 egyptian su	169	6	3.5	190	2	Q9ZBE9	Q9ZBE9 streptomyce
97	106	1	Q9YBT1	Q9YBT1 aeropyrum p	170	6	3.5	192	5	Q9VZG2	Q9VZG2 drosophila
98	108	1	Q9YBT1	Q9YBT1 methanococc	171	6	3.5	193	2	Q93334	Q93334 mycobacteri
99	108	1	Q9YBT1	Q9YBT1 oryza sativ	172	6	3.5	194	1	Q9YB78	Q9YB78 aeropyrum p
100	108	12	Q9LIT5	Q9LIT5 human immun	173	6	3.5	196	13	Q9YH31	Q9YH31 notophtalim
101	108	12	Q9LBM8	Q9LBM8 uncultured	174	6	3.5	198	2	Q9RIW5	Q9RIW5 streptomyce
102	111	4	Q9Y517	Q9Y517 homo sapien	175	6	3.5	199	12	Q9QNZ5	Q9QNZ5 human immun
103	112	2	Q9LBP9	Q9LBP9 chemosynthe	176	6	3.5	201	10	Q9XFY4	Q9XFY4 chlorella p
104	112	2	Q9LBN4	Q9LBN4 uncultured	177	6	3.5	202	2	Q97773	Q97773 mycobacteri
105	112	2	Q9LBM6	Q9LBM6 uncultured	178	6	3.5	202	2	Q9S411	Q9S411 actinobacil
106	112	2	Q9LBM3	Q9LBM3 uncultured	179	6	3.5	203	2	Q97795	Q97795 mycobacteri
107	112	2	Q9LBM2	Q9LBM2 uncultured	180	6	3.5	204	2	Q9RKA5	Q9RKA5 streptomyce
108	112	10	Q9M296	Q9M296 arabidopsis	181	6	3.5	205	2	P95588	P95588 rhodobacter
109	113	2	Q9LBQ1	Q9LBQ1 uncultured	182	6	3.5	206	2	Q9KRH5	Q9KRH5 vibrio chol
110	113	2	Q9LBQ0	Q9LBQ0 uncultured	183	6	3.5	206	10	Q9M282	Q9M282 arabidopsis
111	113	2	Q9LBP8	Q9LBP8 uncultured	184	6	3.5	208	1	Q9UZ35	Q9UZ35 pyrococcus
112	113	10	Q9LDB4	Q9LDB4 arabidopsis	185	6	3.5	210	4	Q9UJ89	Q9UJ89 homo sapien
113	118	10	Q9LK37	Q9LK37 arabidopsis	186	6	3.5	216	10	Q9ZSX5	Q9ZSX5 zea mays (m
114	120	12	Q9LK37	Q9LK37 paramacium	187	6	3.5	217	4	Q9Z589	Q9Z589 homo sapien
115	123	1	Q9YD17	Q9YD17 aeropyrum p	188	6	3.5	218	2	P73776	P73776 synechocyst
116	124	1	Q9YC68	Q9YC68 aeropyrum p	189	6	3.5	218	2	Q9KCB1	Q9KCB1 bacillus ha
117	126	5	Q27011	Q27011 tenebrio mo	190	6	3.5	222	2	Q56620	Q56620 vibrio chol
118	128	2	Q50896	Q50896 myxococcus	191	6	3.5	222	10	Q9ZLVK9	Q9ZLVK9 arabidopsis
119	128	2	Q9WZK5	Q9WZK5 thermotoga	192	6	3.5	222	10	Q9SLW7	Q9SLW7 nicotiana t
120	128	6	Q28195	Q28195 bos taurus	193	6	3.5	223	1	Q9YF37	Q9YF37 aeropyrum p
121	129	4	Q9NT15	Q9NT15 homo sapien	194	6	3.5	225	2	Q34343	Q34343 bacillus su
122	130	1	Q9UW06	Q9UW06 sulfolobus	195	6	3.5	225	6	Q28271	Q28271 canis famil
123	131	4	Q9YF37	Q9YF37 homo sapien	196	6	3.5	226	2	P97177	P97177 rhodobacter
124	132	4	Q9Y3Y7	Q9Y3Y7 lumbicus r	197	6	3.5	227	2	Q9RYM6	Q9RYM6 deinococcus
125	135	8	O47945	O47945 lumbicus r	198	6	3.5	228	1	Q9YEX3	Q9YEX3 aeropyrum p
126	135	10	Q9LGV6	Q9LGV6 oryza sativ	199	6	3.5	229	5	Q9W2Q8	Q9W2Q8 drosophila
127	139	2	P95002	P95002 mycobacteri	200	6	3.5	231	2	Q9PIQ8	Q9PIQ8 campylobact
128	140	10	Q9LPA3	Q9LPA3 arabidopsis	201	6	3.5	232	2	O48821	O48821 lactobacill
129	141	4	Q9NXX4	Q9NXX4 homo sapien	202	6	3.5	232	2	O44138	O44138 lactococcus
130	141	10	Q40437	Q40437 nicotiana s	203	6	3.5	233	2	Q9RLC4	Q9RLC4 pseudomonas
131	144	10	Q23793	Q23793 nicotiana s	204	6	3.5	235	2	Q9Z4Z1	Q9Z4Z1 streptomyce
132	145	2	Q9RYG6	Q9RYG6 deinococcus	205	6	3.5	236	2	Q9JZZ1	Q9JZZ1 neisseria m
133	145	2	Q9RIK1	Q9RIK1 streptococc	206	6	3.5	240	2	O67575	O67575 aquifex ao
134	146	4	Q9UNX1	Q9UNX1 homo sapien	207	6	3.5	242	1	Q9YA61	Q9YA61 aeropyrum p
135	149	1	Q9YB71	Q9YB71 aeropyrum p	208	6	3.5	243	12	O65986	O65986 chino del t
136	149	1	Q9Y9V4	Q9Y9V4 aeropyrum p	209	6	3.5	244	2	Q45845	Q45845 clostridium
137	149	5	Q9NBB4	Q9NBB4 agrotis ips	210	6	3.5	246	11	O88301	O88301 mus musculu
138	151	2	Q53879	Q53879 streptomyce	211	6	3.5	247	5	Q9NDA2	Q9NDA2 trypanosoma
139	151	5	O18728	O18728 caenorhabdi	212	6	3.5	248	2	O68731	O68731 yersinia pe
140	154	2	Q9X8K2	Q9X8K2 streptomyce	213	6	3.5	251	12	O67684	O67684 groundnut r
141	155	2	Q9S2N9	Q9S2N9 streptomyce	214	6	3.5	253	2	Q9WYH9	Q9WYH9 thermotoga
142	155	2	Q9LOE6	Q9LOE6 streptomyce	215	6	3.5	253	5	O18298	O18298 caenorhabdi
143	155	3	Q05669	Q05669 saccharomyc	216	6	3.5	254	2	Q9KGH5	Q9KGH5 bacillus ha
144	156	4	O60888	O60888 homo sapien	217	6	3.5	254	5	Q9V5X5	Q9V5X5 drosophila
145	157	4	Q9NM52	Q9NM52 homo sapien	218	6	3.5	254	10	Q9LWZ9	Q9LWZ9 oryza sativ
146	159	2	Q9RFX4	Q9RFX4 deinococcus	219	6	3.5	256	2	O84256	O84256 chlamydia t
147	160	4	Q9F1B9	Q9F1B9 homo sapien	220	6	3.5	257	5	Q26653	Q26653 sarcophaga
148	161	12	Q04531	Q04531 pseudorabie	221	6	3.5	258	12	O55501	O55501 onion yello
149	162	11	P70598	P70598 rattus norv	222	6	3.5	258	12	O55502	O55502 onion yello
150	164	2	O51260	O51260 borrelia bu	223	6	3.5	259	2	O51977	O51977 pseudomonas
151	167	3	O01439	O01439 netelia hae	224	6	3.5	260	3	Q9U079	Q9U079 schizosacch
152	168	5	Q06065	Q06065 mycobacteri	225	6	3.5	261	4	Q9UP16	Q9UP16 homo sapien
153	168	5	Q9NNH5	Q9NNH5 leishmania	226	6	3.5	262	2	Q49810	Q49810 mycobacteri
154	168	12	O84423	O84423 paramacium	227	6	3.5	263	6	Q9MXX2	Q9MXX2 sus scrofa
155	169	2	O07429	O07429 mycobacteri	228	6	3.5	263	10	Q9LGS2	Q9LGS2 oryza sativ
156	169	2	Q9KOP7	Q9KOP7 neisseria m	229	6	3.5	265	2	Q52682	Q52682 rhodobacter
157	169	10	Q9LWD1	Q9LWD1 oryza sativ	230	6	3.5	265	2	O68089	O68089 rhodobacter
158	171	2	Q9JVS5	Q9JVS5 neisseria m	231	6	3.5	265	2	Q9X8N6	Q9X8N6 streptomyce
159	174	5	O02082	O02082 caenorhabdi	232	6	3.5	267	12	O42027	O42027 welsh onlon
160	175	2	P73596	P73596 synechocyst	233	6	3.5	268	2	Q02303	Q02303 pseudomonas
161	175	10	Q9LTY8	Q9LTY8 arabidopsis	234	6	3.5	269	2	Q9RYR9	Q9RYR9 deinococcus
162	176	10	Q9LIJ0	Q9LIJ0 arabidopsis	235	6	3.5	269	5	Q9UI68	Q9UI68 leishmania
163	177	12	O41996	O41996 maize rayad	236	6	3.5	271	2	Q9S218	Q9S218 streptomyce
164	179	11	Q9JWK9	Q9JWK9 mus musculu	237	6	3.5	273	2	O53624	O53624 mycobacteri
165	180	4	Q9Y479	Q9Y479 homo sapien	238	6	3.5	275	6	Q9N2D1	Q9N2D1 sus scrofa

239	6	3.5	276	2	Q9RYH3	Q9ryh3 deinococcus	312	6	3.5	342	5	O61580	O61580 drosophila
240	6	3.5	279	5	Q19882	Q19882 caenorhabdi	313	6	3.5	345	12	Q9YWH5	Q9ywh5 influenza b
241	6	3.5	281	5	Q23120	Q23120 caenorhabdi	314	6	3.5	345	12	Q9YWH4	Q9ywh4 influenza b
242	6	3.5	283	4	Q92956	Q92956 homo sapien	315	6	3.5	345	12	Q9YWH3	Q9ywh3 influenza b
243	6	3.5	283	4	Q9UM65	Q9um65 homo sapien	316	6	3.5	345	12	Q9YWH2	Q9ywh2 influenza b
244	6	3.5	283	6	Q9XS28	Q9xs28 cercopitheci	317	6	3.5	345	12	Q9YIL8	Q9yil8 influenza b
245	6	3.5	288	4	Q92999	Q92999 homo sapien	318	6	3.5	347	2	Q9RD12	Q9rd12 streptomyce
246	6	3.5	291	4	Q92590	Q92590 homo sapien	319	6	3.5	347	5	Q9XX84	Q9xx84 caenorhabdi
247	6	3.5	292	3	Q9Y7L8	Q9y7l8 schizosacch	320	6	3.5	349	1	O28037	O28037 archaeoglob
248	6	3.5	292	12	Q9WP55	Q9wf55 bean calico	321	6	3.5	350	5	O18323	O18323 caenorhabdi
249	6	3.5	293	5	Q9N788	Q9n788 leishmania	322	6	3.5	350	8	Q9TD20	Q9td20 neofundulus
250	6	3.5	293	12	O56374	O56374 sida golden	323	6	3.5	353	10	Q9M4Q7	Q9m4q7 ricinus com
251	6	3.5	293	12	O71208	O71208 potatato yell	324	6	3.5	355	2	Q9L1P4	Q9l1p4 streptomyce
252	6	3.5	293	12	O96615	O96615 abutilon mo	325	6	3.5	355	2	Q9K950	Q9k950 bacillus ha
253	6	3.5	293	12	O36626	O36626 taino tomat	326	6	3.5	359	2	O05139	O05139 agrobacteri
254	6	3.5	293	12	O41444	O41444 tomatato mott	327	6	3.5	359	8	Q37749	Q37749 cephalospor
255	6	3.5	293	12	O41445	O41445 tomatato mott	328	6	3.5	360	2	Q9X613	Q9x613 chlorobium
256	6	3.5	293	12	O36456	O36456 chino del t	329	6	3.5	360	2	O53282	O53282 mycobacteri
257	6	3.5	293	12	O01734	O01734 bean golden	330	6	3.5	360	2	Q9R692	Q9r692 agrobacteri
258	6	3.5	293	12	O67563	O67563 bean dwarf	331	6	3.5	360	10	O9LFC0	O9lfc0 arabidopsis
259	6	3.5	293	12	O67586	O67586 bean golden	332	6	3.5	361	4	Q9NPQ2	Q9npq2 homo sapien
260	6	3.5	293	12	O41334	O41334 tomatato leaf	333	6	3.5	362	2	Q9RD10	Q9rd10 streptomyce
261	6	3.5	293	12	P89118	P89118 sida golden	334	6	3.5	362	12	Q9IJY0	Q9ijy0 influenza b
262	6	3.5	293	12	P89120	P89120 sida golden	335	6	3.5	362	12	Q9IJX9	Q9ijx9 influenza b
263	6	3.5	293	12	O67573	O67573 bean golden	336	6	3.5	367	2	Q9ZNB1	Q9znb1 streptomyce
264	6	3.5	293	12	Q9YL71	Q9yl71 tomatato leaf	337	6	3.5	367	2	Q9KLY7	Q9kly7 vibrio chol
265	6	3.5	293	12	P89121	P89121 sida golden	338	6	3.5	370	4	O75194	O75194 homo sapien
266	6	3.5	293	12	Q9QGH1	Q9qgh1 bean golden	339	6	3.5	372	4	Q9NZR6	Q9nzh6 homo sapien
267	6	3.5	293	12	Q9Q9A3	Q9q9a3 dicliptera	340	6	3.5	372	11	Q9QZP4	Q9qzp4 mus musculus
268	6	3.5	295	11	Q9WTJ6	Q9wtj6 mus musculus	341	6	3.5	372	12	O11294	O11294 molluscum c
269	6	3.5	295	12	O91203	O91203 havana toma	342	6	3.5	373	10	Q9L776	Q9l776 arabidopsis
270	6	3.5	298	2	Q9ZJZ2	Q9zjz2 helicobacte	343	6	3.5	374	1	Q9VCM8	Q9vcm8 aeropyrum p
271	6	3.5	298	2	Q9RHL0	Q9rhl0 rhodopsendo	344	6	3.5	374	2	O9ISV1	O9isv1 neisseria m
272	6	3.5	300	2	Q9Z1I9	Q9z1i9 streptomyce	345	6	3.5	375	2	O50455	O50455 mycobacteri
273	6	3.5	300	2	Q9XBI4	Q9xbi4 mycobacteri	346	6	3.5	375	2	Q9WX77	Q9wx77 thermus aqu
274	6	3.5	300	2	Q9L030	Q9l030 streptomyce	347	6	3.5	375	2	Q9KZL8	Q9kzl8 streptomyce
275	6	3.5	301	4	Q9UFL0	Q9uf10 homo sapien	348	6	3.5	376	2	O87140	O87140 vibrio chol
276	6	3.5	301	4	Q9UEW7	Q9uew7 homo sapien	349	6	3.5	376	2	Q9ZD59	Q9zdz9 rickettsia
277	6	3.5	301	11	Q9Z2U5	Q9z2u5 mus musculus	350	6	3.5	377	2	O34236	O34236 vibrio chol
278	6	3.5	303	2	O68482	O68482 achromobact	351	6	3.5	377	2	Q9K128	Q9k128 neisseria m
279	6	3.5	303	2	Q9K3X6	Q9k3x6 streptomyce	352	6	3.5	378	12	O65822	O65822 bovine herp
280	6	3.5	305	10	Q9SIG0	Q9sig0 arabidopsis	353	6	3.5	379	1	O58782	O58782 methanococc
281	6	3.5	306	2	Q9KTI3	Q9kti3 vibrio chol	354	6	3.5	379	8	Q9TEF9	Q9tef9 cavia porce
282	6	3.5	307	2	Q9PEV1	Q9pev1 xyella fas	355	6	3.5	381	5	O76183	O76183 amphidinium
283	6	3.5	309	11	Q9QVZ0	Q9qvz0 mus sp. shb	356	6	3.5	382	8	Q9Z2R0	Q9z2r0 cyanidiosch
284	6	3.5	311	2	Q9RQK0	Q9rk0 caulobacter	357	6	3.5	382	10	Q9ZPF4	Q9zpf4 arabidopsis
285	6	3.5	312	1	O98550	O98550 pyrococcus	358	6	3.5	383	2	O11064	O11064 mycobacteri
286	6	3.5	312	4	Q92977	Q92977 homo sapien	359	6	3.5	383	10	O50063	O50063 arabidopsis
287	6	3.5	312	4	Q99620	Q99620 homo sapien	360	6	3.5	384	2	Q9KY59	Q9ky59 streptomyce
288	6	3.5	313	2	O87265	O87265 pseudomonas	361	6	3.5	384	2	O9JWJ4	O9jwj4 neisseria m
289	6	3.5	316	2	Q9Z5B8	Q9z5b8 streptomyce	362	6	3.5	384	5	O9N4J1	O9n4j1 caenorhabdi
290	6	3.5	317	1	Q9YD93	Q9yd93 aeropyrum p	363	6	3.5	385	2	Q9X9K7	Q9x9k7 thermus aqu
291	6	3.5	320	2	Q9RHC9	Q9rhc9 pseudomonas	364	6	3.5	385	5	O17909	O17909 caenorhabdi
292	6	3.5	320	2	Q9KWT2	Q9kwt2 vibrio algi	365	6	3.5	386	2	Q9K7W8	Q9k7w8 vibrio chol
293	6	3.5	320	2	Q9KRG1	Q9kr91 vibrio chol	366	6	3.5	387	2	Q9RNM6	Q9rnm6 zymomonas m
294	6	3.5	321	10	O23891	O23891 oryza sativ	367	6	3.5	387	10	Q9ZNO5	Q9zn05 arabidopsis
295	6	3.5	321	11	O70341	O70341 mus musculus	368	6	3.5	388	12	O65360	O65360 orgyia pseu
296	6	3.5	322	13	Q9P7B4	Q9ptb4 brachydanio	369	6	3.5	388	13	Q9YH95	Q9yh95 xenopus lae
297	6	3.5	323	4	Q9Y405	Q9y405 homo sapien	370	6	3.5	389	4	Q9NZJ7	Q9nzh7 homo sapien
298	6	3.5	325	5	O16086	O16086 eumesocampa	371	6	3.5	389	10	Q9M929	Q9m929 arabidopsis
299	6	3.5	330	4	O16086	O16086 homo sapien	372	6	3.5	389	11	O9QZ45	O9qz45 mus musculus
300	6	3.5	332	2	Q9RW60	Q9rw60 deinococcus	373	6	3.5	392	6	O28414	O28414 felis silve
301	6	3.5	332	5	O9N451	Q9n451 caenorhabdi	374	6	3.5	392	12	O92811	O92811 bovine leuk
302	6	3.5	333	10	Q9SQP0	Q9sqp0 arabidopsis	375	6	3.5	393	2	Q9LC16	Q9lc16 streptomyce
303	6	3.5	333	10	Q9LM58	Q9lm58 arabidopsis	376	6	3.5	393	2	O9LC15	O9lc15 streptomyce
304	6	3.5	335	2	Q9KES7	Q9kes7 bacillus ha	377	6	3.5	393	2	O9LC14	O9lc14 streptomyce
305	6	3.5	336	2	Q9K3T6	Q9k3t6 streptomyce	378	6	3.5	393	2	O9LC13	O9lc13 streptomyce
306	6	3.5	336	11	Q9JMW0	Q9jmw0 rattus norv	379	6	3.5	393	12	O85490	O85490 bovine leuk
307	6	3.5	338	2	Q9RT55	Q9rt55 deinococcus	380	6	3.5	396	1	Q9YD77	Q9ydz7 aeropyrum p
308	6	3.5	338	10	Q9LSG7	Q9ls97 arabidopsis	381	6	3.5	396	2	Q9RDF1	Q9rdf1 streptomyce
309	6	3.5	338	11	O88880	O88880 mus musculus	382	6	3.5	397	2	Q9RJR4	Q9rjr4 streptomyce
310	6	3.5	340	10	Q9MAG5	Q9mag5 arabidopsis	383	6	3.5	398	2	O9ZAL3	O9zal3 pseudomonas
311	6	3.5	340	11	O35125	O35125 mus musculus	384	6	3.5	399	4	Q9UJU7	Q9uju7 homo sapien

385	6	3.5	401	11	Q9QYK5	Q9qyk5 mus musculus	458	6	3.5	465	2	Q9KSB6	Q9ksb6 vibrio chol
386	6	3.5	406	12	Q66616	Q66616 equine herp	459	6	3.5	468	2	Q9KS68	Q9ks68 vibrio chol
387	6	3.5	406	4	Q9NVA8	Q9nva8 homo sapien	460	6	3.5	469	1	Q9YFCA	Q9yfc4 aeropyrum p
388	6	3.5	406	5	Q22192	Q22192 caenorhabdi	461	6	3.5	469	11	Q63721	Q63721 rattus norv
389	6	3.5	408	2	Q9L0L5	Q9l0l5 streptomyce	462	6	3.5	470	5	Q9XUY1	Q9xuy1 caenorhabdi
390	6	3.5	409	2	Q9K9Q3	Q9k9q3 bacillus ha	463	6	3.5	472	4	Q9Y6J1	Q9y6j1 homo sapien
391	6	3.5	410	4	Q60243	Q60243 homo sapien	464	6	3.5	473	2	Q9X5Z0	Q9x5z0 thibacillu
392	6	3.5	410	11	Q63720	Q63720 rattus norv	465	6	3.5	473	5	Q9Y0F2	Q9y0f2 toxoplasma
393	6	3.5	410	11	Q9JMJ2	Q9jmj2 mus musculus	466	6	3.5	475	10	Q9SJ64	Q9sj64 arabidopsis
394	6	3.5	411	2	Q69522	Q69522 mycobacteri	467	6	3.5	476	2	Q9KRNO	Q9krn0 vibrio chol
395	6	3.5	412	2	Q57031	Q57031 rhodococcus	468	6	3.5	476	5	Q9W052	Q9w052 drosophila
396	6	3.5	413	2	Q69679	Q69679 mycobacteri	469	6	3.5	478	2	Q9LBB7	Q9lbb7 gluconobact
397	6	3.5	413	10	Q9LXSO	Q9lxs0 arabidopsis	470	6	3.5	480	2	Q9X5B7	Q9x5b7 clostridium
398	6	3.5	413	10	Q9L126	Q9l126 oryza sativ	471	6	3.5	480	2	Q53244	Q53244 mycobacteri
399	6	3.5	416	2	Q9S2S0	Q9s2s0 streptomyce	472	6	3.5	482	2	Q9KFP6	Q9kfp6 bacillus ha
400	6	3.5	416	4	Q94781	Q94781 homo sapien	473	6	3.5	482	5	Q09937	Q09937 caenorhabdi
401	6	3.5	416	4	Q90QV4	Q90qv4 homo sapien	474	6	3.5	484	3	Q12234	Q12234 saccharomyc
402	6	3.5	417	1	Q57907	Q57907 pyrococcus	475	6	3.5	485	2	Q06783	Q06783 streptomyce
403	6	3.5	417	2	Q9RPA9	Q9rpa9 streptomyce	476	6	3.5	485	10	Q9ZQ07	Q9zq07 arabidopsis
404	6	3.5	417	10	Q9SFE3	Q9sfe3 arabidopsis	477	6	3.5	488	2	Q9K0U2	Q9k0u2 neisseria m
405	6	3.5	418	10	Q9W1Y1	Q9w1y1 arabidopsis	478	6	3.5	488	2	Q9JT26	Q9jtt26 neisseria m
406	6	3.5	420	3	P78892	P78892 schizosacch	479	6	3.5	490	5	Q9XDX1	Q9xxd1 caenorhabdi
407	6	3.5	420	11	Q9QX71	Q9qx71 rattus norv	480	6	3.5	494	2	Q9RWF6	Q9rwf6 deinococcus
408	6	3.5	421	12	Q98320	Q98320 molluscum c	481	6	3.5	495	2	Q51281	Q51281 neisseria m
409	6	3.5	421	12	Q9WSP0	Q9wsp0 human immun	482	6	3.5	497	5	Q9VNUJ1	Q9vnuj1 drosophila
410	6	3.5	422	2	Q9WXB4	Q9wxb4 acidiphiliu	483	6	3.5	498	10	Q39483	Q39483 calocedrus
411	6	3.5	422	12	Q9WSP7	Q9wsp7 human immun	484	6	3.5	500	4	Q9UHH4	Q9uhj4 homo sapien
412	6	3.5	423	2	Q48375	Q48375 kibdelospor	485	6	3.5	501	2	Q9RVX7	Q9rvx7 deinococcus
413	6	3.5	423	2	Q9L551	Q9l551 streptomyce	486	6	3.5	501	5	Q9W0I4	Q9w0i4 drosophila
414	6	3.5	428	2	Q9KXU8	Q9kxu8 streptomyce	487	6	3.5	503	11	P97557	P97557 rattus norv
415	6	3.5	430	2	P96664	P96664 bacillus su	488	6	3.5	503	12	Q74827	Q74827 human immun
416	6	3.5	430	6	Q9N092	Q9n092 macaca fasc	489	6	3.5	504	11	Q60565	Q60565 mesocricetu
417	6	3.5	431	1	Q57908	Q57908 pyrococcus	490	6	3.5	504	11	Q9J188	Q9j188 rattus norv
418	6	3.5	432	5	Q9W2N5	Q9w2n5 drosophila	491	6	3.5	504	11	Q9JHE5	Q9jhe5 rattus norv
419	6	3.5	433	10	Q9SNT6	Q9snt6 oryza sativ	492	6	3.5	505	10	Q22878	Q22878 arabidopsis
420	6	3.5	434	2	Q9L0K6	Q9l0k6 streptomyce	493	6	3.5	506	5	Q7598	Q7598 caenorhabdi
421	6	3.5	434	3	Q02933	Q02933 saccharomyc	494	6	3.5	507	2	Q9S1E5	Q9s1e5 wolnelli s
422	6	3.5	434	4	Q14829	Q14829 homo sapien	495	6	3.5	507	13	Q57439	Q57439 xenopus lae
423	6	3.5	435	1	Q9UYB8	Q9uyb8 pyrococcus	496	6	3.5	507	13	Q13020	Q13020 xenopus lae
424	6	3.5	435	5	Q9W3X3	Q9w3x3 drosophila	497	6	3.5	509	12	Q9YX49	Q9yx49 human immun
425	6	3.5	435	10	Q9SLP3	Q9slp3 cucumis sat	498	6	3.5	510	2	Q9PY7	Q9py7 xylella fas
426	6	3.5	436	2	Q9X6N7	Q9x6n7 rhizobium l	499	6	3.5	512	2	Q9KEF8	Q9kef8 bacillus ha
427	6	3.5	437	10	Q9L2P7	Q9l2p7 arabidopsis	500	6	3.5	513	2	Q9PLG8	Q9plg8 chlamydia m
428	6	3.5	438	10	Q82244	Q82244 arabidopsis	501	6	3.5	513	10	Q9M8V5	Q9m8v5 arabidopsis
429	6	3.5	439	2	Q9XCX0	Q9xcx0 rhodobacter	502	6	3.5	514	2	Q9Z4P4	Q9z4p4 sulfurospir
430	6	3.5	439	2	P71616	P71616 mycobacteri	503	6	3.5	515	2	Q9RRK1	Q9rrk1 deinococcus
431	6	3.5	440	2	Q34936	Q34936 bacillus su	504	6	3.5	516	10	Q9SSP5	Q9ssp5 arabidopsis
432	6	3.5	440	2	Q9RXX0	Q9rx0 streptomyce	505	6	3.5	516	10	Q9ZWT3	Q9zwt3 arabidopsis
433	6	3.5	443	2	Q50205	Q50205 xanthobacte	506	6	3.5	517	5	Q9V7M9	Q9v7m9 drosophila
434	6	3.5	443	2	Q53919	Q53919 mycobacteri	507	6	3.5	519	8	Q9MT28	Q9mt28 solanum tub
435	6	3.5	445	2	Q9LIA8	Q9lia8 streptomyce	508	6	3.5	520	10	Q49522	Q49522 arabidopsis
436	6	3.5	445	2	Q31777	Q31777 mycobacteri	509	6	3.5	523	2	Q56839	Q56839 xanthobacte
437	6	3.5	446	2	Q9KVM0	Q9kvm0 streptomyce	510	6	3.5	524	4	Q9NSU1	Q9nsul homo sapien
438	6	3.5	447	2	Q9L2G5	Q9l2g5 streptomyce	511	6	3.5	524	5	Q9VLY6	Q9vly6 drosophila
439	6	3.5	447	10	Q9SRM7	Q9srm7 arabidopsis	512	6	3.5	524	10	Q39722	Q39722 ephedra ger
440	6	3.5	448	4	Q9NM00	Q9nm00 homo sapien	513	6	3.5	527	12	Q72605	Q72605 human immun
441	6	3.5	449	5	Q26978	Q26978 trypanosoma	514	6	3.5	528	2	P71969	P71969 mycobacteri
442	6	3.5	450	2	Q9ZIV1	Q9ziv1 megasphaera	515	6	3.5	528	5	P91256	P91256 caenorhabdi
443	6	3.5	450	11	Q9JJJ9	Q9jjj9 mus musculus	516	6	3.5	528	5	Q9VJM4	Q9vjm4 drosophila
444	6	3.5	451	2	Q9KNN2	Q9knn2 vibrio chol	517	6	3.5	528	13	Q93440	Q93440 gallus gall
445	6	3.5	454	5	Q9VM69	Q9vm69 drosophila	518	6	3.5	529	2	Q9RVQ7	Q9rvq7 deinococcus
446	6	3.5	455	11	Q9JJJ8	Q9jjj8 mus musculus	519	6	3.5	532	5	Q96671	Q96671 drosophila
447	6	3.5	456	10	Q9LU25	Q9lu25 arabidopsis	520	6	3.5	534	5	Q9NKK69	Q9nkk69 drosophila
448	6	3.5	457	11	Q9JJK0	Q9jjk0 mus musculus	521	6	3.5	536	2	Q9RYP0	Q9ryp0 deinococcus
449	6	3.5	457	10	Q9ZUE7	Q9zue7 arabidopsis	522	6	3.5	537	3	O42801	O42801 aspergillus
450	6	3.5	458	2	Q9RU65	Q9ru65 deinococcus	523	6	3.5	539	2	O68036	O68036 rhodobacter
451	6	3.5	459	2	O86622	O86622 streptomyce	524	6	3.5	539	11	Q60664	Q60664 mus musculus
452	6	3.5	461	10	Q9LYT6	Q9lyt6 arabidopsis	525	6	3.5	540	4	Q9NPX0	Q9npx0 homo sapien
453	6	3.5	461	11	Q9JJJ7	Q9jjj7 mus musculus	526	6	3.5	540	5	Q9W053	Q9w053 drosophila
454	6	3.5	462	2	Q51663	Q51663 paracoccus	527	6	3.5	546	2	P72175	P72175 pseudomonas
455	6	3.5	462	4	Q9P2G5	Q9p2g5 homo sapien	528	6	3.5	546	8	Q35989	Q35989 trypanoplas
456	6	3.5	463	10	Q22145	Q22145 arabidopsis	529	6	3.5	547	2	Q9RFM9	Q9rfm9 pseudomonas
457	6	3.5	464	2	Q48419	Q48419 klebsiella	530	6	3.5	548	8	O63543	O63543 bodo saltan

531	6	3.5	555	4	Q16764	Q16764 homo sapien	604	6	3.5	671	4	O95559	O95559 homo sapien
532	6	3.5	555	4	Q9P103	Q9P103 homo sapien	605	6	3.5	676	10	O9LYA8	O9LYA8 arabidopsis
533	6	3.5	555	4	Q9N077	Q9N077 homo sapien	606	6	3.5	683	6	O9N091	Q9N091 macaca fasc
534	6	3.5	555	5	Q9U0V4	Q9U0V4 leishmania	607	6	3.5	684	2	Q44055	Q44055 aeromonas h
535	6	3.5	557	10	Q9M3C5	Q9M3C5 arabidopsis	608	6	3.5	684	2	Q9RVA3	Q9RVA3 deinococcus
536	6	3.5	559	10	Q9SDL3	Q9SDL3 apium grave	609	6	3.5	685	10	Q9ZSB2	Q9ZSB2 arabidopsis
537	6	3.5	560	4	Q9UMN0	Q9UMN0 homo sapien	610	6	3.5	685	10	P91352	P91352 caenorhabdi
538	6	3.5	564	2	O53539	O53539 mycobacteri	611	6	3.5	687	5	P91352	Q9K186 caenorhabdi
539	6	3.5	567	2	Q9KYR6	Q9KYR6 streptomyc	612	6	3.5	690	2	Q9K186	Q9K186 vibrio chol
540	6	3.5	569	5	Q9U0X1	Q9U0X1 leishmania	613	6	3.5	697	5	O61121	O61121 crithidia f
541	6	3.5	569	11	Q9QX11	Q9QX11 mus musculu	614	6	3.5	699	10	Q9ZU69	Q9ZU69 arabidopsis
542	6	3.5	570	2	Q9KQM3	Q9KQM3 vibrio chol	615	6	3.5	700	5	Q9V613	Q9V613 drosophila
543	6	3.5	571	10	Q9LE01	Q9LE01 oryza sativ	616	6	3.5	702	5	Q9VU16	Q9VU16 drosophila
544	6	3.5	571	12	Q9WJP3	Q9WJP3 bovine leuk	617	6	3.5	702	12	P89440	P89440 herpes simp
545	6	3.5	572	11	O61090	O61090 mus musculu	618	6	3.5	708	10	Q9LEN6	Q9LEN6 cicler ariet
546	6	3.5	574	2	O87312	O87312 mycobacteri	619	6	3.5	717	5	O9VKW5	O9VKW5 drosophila
547	6	3.5	575	2	Q9ZB57	Q9ZB57 proteus mir	620	6	3.5	718	10	Q9SM79	Q9SM79 oryza sativ
548	6	3.5	577	5	Q16976	Q16976 aplysia cal	621	6	3.5	718	10	Q9SVT2	Q9SVT2 arabidopsis
549	6	3.5	579	2	Q9KAB5	Q9KAB5 bacillus ha	622	6	3.5	719	2	Q9L5T4	Q9L5T4 salmonella
550	6	3.5	579	4	O15331	O15331 homo sapien	623	6	3.5	721	5	Q9NAB9	Q9NAB9 caenorhabdi
551	6	3.5	580	2	Q9K3W7	Q9K3W7 streptomyc	624	6	3.5	726	4	Q43506	Q43506 homo sapien
552	6	3.5	581	1	O29674	O29674 archaeoglob	625	6	3.5	726	4	Q9UKT9	Q9UKT9 homo sapien
553	6	3.5	581	5	Q23969	Q23969 drosophila	626	6	3.5	730	11	O88735	O88735 mus musculu
554	6	3.5	585	3	O74928	O74928 schizosacch	627	6	3.5	731	4	Q9NUH2	Q9NUH2 homo sapien
555	6	3.5	586	3	Q07824	Q07824 saccharomyc	628	6	3.5	733	2	Q9X7Y4	Q9X7Y4 streptomyc
556	6	3.5	589	2	Q9K7H7	Q9K7H7 bacillus ha	629	6	3.5	735	11	Q9JIG4	Q9JIG4 mus musculu
557	6	3.5	590	4	O00456	O00456 homo sapien	630	6	3.5	738	5	Q9XWV5	Q9XWV5 caenorhabdi
558	6	3.5	590	10	O9MAH4	O9MAH4 arabidopsis	631	6	3.5	738	11	Q9Z213	Q9Z213 rattus norv
559	6	3.5	592	4	Q9NFP1	Q9NFP1 homo sapien	632	6	3.5	739	4	Q9NS39	Q9NS39 homo sapien
560	6	3.5	593	2	O83470	O83470 treponema p	633	6	3.5	740	2	Q9PG63	Q9PG63 xylella fas
561	6	3.5	593	5	Q9XTY6	Q9XTY6 caenorhabdi	634	6	3.5	746	10	Q9STQ2	Q9STQ2 arabidopsis
562	6	3.5	594	5	O04579	O04579 drosophila	635	6	3.5	747	10	Q9ZSB1	Q9ZSB1 arabidopsis
563	6	3.5	596	5	Q9V623	Q9V623 drosophila	636	6	3.5	748	2	Q9K242	Q9K242 streptomyc
564	6	3.5	599	5	Q24629	Q24629 drosophila	637	6	3.5	749	2	Q9RY43	Q9RY43 deinococcus
565	6	3.5	603	10	Q9LVE3	Q9LVE3 arabidopsis	638	6	3.5	751	2	Q9ZAA5	Q9ZAA5 streptomyc
566	6	3.5	605	2	Q9S0T2	Q9S0T2 escherichia	639	6	3.5	751	10	Q9ZSB3	Q9ZSB3 arabidopsis
567	6	3.5	605	2	Q9PMV9	Q9PMV9 campylobact	640	6	3.5	751	10	Q9ZT72	Q9ZT72 arabidopsis
568	6	3.5	607	2	Q9KY63	Q9KY63 streptomyc	641	6	3.5	753	2	Q9RD19	Q9RD19 streptomyc
569	6	3.5	607	2	Q9PMV8	Q9PMV8 campylobact	642	6	3.5	753	10	O9LWY1	O9LWY1 arabidopsis
570	6	3.5	608	13	Q9W7F2	Q9W7F2 xenopus lae	643	6	3.5	755	10	O9LYH9	O9LYH9 arabidopsis
571	6	3.5	609	5	Q94834	Q94834 trichomonas	644	6	3.5	755	11	Q9R0G6	Q9R0G6 mus musculu
572	6	3.5	609	5	Q9NDR0	Q9NDR0 ciona intes	645	6	3.5	756	10	Q9ZSB0	Q9ZSB0 arabidopsis
573	6	3.5	610	5	Q9VUB4	Q9VUB4 drosophila	646	6	3.5	757	2	Q9K1B0	Q9K1B0 porphyromon
574	6	3.5	611	12	P89885	P89885 onion yello	647	6	3.5	760	10	O80866	O80866 arabidopsis
575	6	3.5	612	11	Q9Z2N3	Q9Z2N3 mus musculu	648	6	3.5	763	10	O9WAP6	O9WAP6 arabidopsis
576	6	3.5	614	4	Q03989	Q03989 homo sapien	649	6	3.5	765	10	O9SZY2	O9SZY2 arabidopsis
577	6	3.5	614	13	Q91134	Q91134 naja naja (650	6	3.5	769	10	Q9SVT3	Q9SVT3 arabidopsis
578	6	3.5	615	2	O86143	O86143 corynebacte	651	6	3.5	771	2	O06359	O06359 mycobacteri
579	6	3.5	615	5	O9N9Q6	Q9N9Q6 leishmania	652	6	3.5	772	10	Q9SVT4	Q9SVT4 arabidopsis
580	6	3.5	616	10	Q9LWU5	Q9LWU5 arabidopsis	653	6	3.5	774	10	Q9ZSA9	Q9ZSA9 streptomyc
581	6	3.5	617	10	O9LWU9	Q9LWU9 oryza sativ	654	6	3.5	774	10	O9WAP7	Q9WAP7 arabidopsis
582	6	3.5	619	5	O76307	O76307 trichomonas	655	6	3.5	774	4	O60347	O60347 homo sapien
583	6	3.5	622	4	O14794	O14794 homo sapien	656	6	3.5	775	4	O9SZY3	O9SZY3 arabidopsis
584	6	3.5	622	3	Q9Y786	Q9Y786 magnaporthe	657	6	3.5	775	10	Q9WAK9	Q9WAK9 human rotav
585	6	3.5	629	3	Q9Y785	Q9Y785 magnaporthe	658	6	3.5	776	2	O86516	O86516 streptomyc
586	6	3.5	631	3	Q9Y784	Q9Y784 magnaporthe	659	6	3.5	777	3	Q9USN3	Q9USN3 schizosacch
587	6	3.5	637	10	Q03678	Q03678 hordeum vul	660	6	3.5	777	10	O9WAP5	O9WAP5 arabidopsis
588	6	3.5	641	10	Q9ZT73	Q9ZT73 arabidopsis	661	6	3.5	783	2	Q9XAL1	Q9XAL1 streptomyc
589	6	3.5	642	5	Q9N2T9	Q9N2T9 caenorhabdi	662	6	3.5	788	2	Q9PFA1	Q9PFA1 xylella fas
590	6	3.5	646	2	O06981	O06981 bacillus su	663	6	3.5	789	12	Q9YXE5	Q9YXE5 sphaeropsis
591	6	3.5	646	2	O67197	O67197 aquifex aeo	664	6	3.5	798	10	Q9LEX8	Q9LEX8 arabidopsis
592	6	3.5	647	2	Q9WZE0	Q9WZE0 thermotoga	665	6	3.5	802	5	Q9VJ74	Q9VJ74 drosophila
593	6	3.5	649	10	Q9SX87	Q9SX87 arabidopsis	666	6	3.5	803	4	Q9UNY1	Q9UNY1 homo sapien
594	6	3.5	649	10	Q9W549	Q9W549 cucumis sac	667	6	3.5	803	10	Q9T0B5	Q9T0B5 arabidopsis
595	6	3.5	652	11	O54858	O54858 rattus norv	668	6	3.5	816	5	Q9V6C1	Q9V6C1 drosophila
596	6	3.5	652	11	O54859	O54859 rattus norv	669	6	3.5	817	4	O75420	O75420 homo sapien
597	6	3.5	653	10	Q9SKK6	Q9SKK6 arabidopsis	670	6	3.5	826	10	O64608	O64608 arabidopsis
598	6	3.5	654	2	P74664	P74664 synecocyst	671	6	3.5	830	10	O9SUG2	O9SUG2 arabidopsis
599	6	3.5	655	5	Q94832	Q94832 trichomonas	672	6	3.5	835	2	P74210	P74210 synecocyst
600	6	3.5	656	2	Q9KK51	Q9KK51 vibrio chol	673	6	3.5	836	12	Q91DV2	Q91DV2 human immun
601	6	3.5	656	10	Q9SGA1	Q9SGA1 arabidopsis	674	6	3.5	839	5	Q9Y103	Q9Y103 drosophila
602	6	3.5	659	4	O96026	O96026 homo sapien	675	6	3.5	839	5	Q9NHN7	Q9NHN7 drosophila
603	6	3.5	660	12	O36613	O36613 swine hepat	676	6	3.5	840	12	Q9QSU2	Q9QSU2 human immun

677	6	3.5	842	4	Q9NP58	Q9np58 homo sapien	750	6	3.5	1024	5	Q9VM97	Q9vm97 drosophila
678	6	3.5	845	2	P74119	P74119 synchocyst	751	6	3.5	1038	3	Q9P4C6	Q9p4c6 pichia past
679	6	3.5	848	2	Q46837	Q46837 escherichia	752	6	3.5	1039	3	Q9N974	Q9n974 leishmania
680	6	3.5	850	12	Q9IW37	Q9iw37 human immun	753	6	3.5	1044	3	Q94751	Q94751 schizosacch
681	6	3.5	854	2	Q84288	Q84288 chlamydia t	754	6	3.5	1052	4	Q9Y4C4	Q9y4c4 homo sapien
682	6	3.5	855	6	Q9TU00	Q9tu00 bos taurus	755	6	3.5	1059	5	Q9VVC6	Q9vvc6 drosophila
683	6	3.5	855	12	Q72619	Q72619 human immun	756	6	3.5	1071	10	Q9MR97	Q9mr97 arabidopsis
684	6	3.5	857	4	Q9UHP2	Q9uhp2 homo sapien	757	6	3.5	1083	2	Q9RPH6	Q9rph6 mycobacteri
685	6	3.5	858	12	Q75114	Q75114 human immun	758	6	3.5	1091	2	Q9Z573	Q9z573 streptomyce
686	6	3.5	859	6	Q9N1E9	Q9n1e9 bos taurus	759	6	3.5	1095	10	Q9LUZ5	Q9luz5 arabidopsis
687	6	3.5	859	12	Q65143	Q65143 african swi	760	6	3.5	1097	5	Q9U2T9	Q9u2t9 caenorhabdi
688	6	3.5	859	12	Q9QRX1	Q9qrx1 human immun	761	6	3.5	1099	5	Q9VMS5	Q9vms5 drosophila
689	6	3.5	861	10	Q9SLS3	Q9sls3 nicotiana t	762	6	3.5	1110	5	P92160	P92160 caenorhabdi
690	6	3.5	861	12	Q83472	Q83472 southern be	763	6	3.5	1111	4	Q75155	Q75155 homo sapien
691	6	3.5	861	12	Q90599	Q90599 human immun	764	6	3.5	1134	1	Q30078	Q30078 archaeoglob
692	6	3.5	862	12	Q9WQ12	Q9wq12 human immun	765	6	3.5	1155	2	Q9KW2	Q9kw2 vibrio chol
693	6	3.5	863	12	Q9IV26	Q9iv26 human immun	766	6	3.5	1160	12	Q9WP26	Q9wp26 bovine vira
694	6	3.5	865	12	Q9IV12	Q9iv12 human immun	767	6	3.5	1161	10	Q9LQD1	Q9lqd1 arabidopsis
695	6	3.5	870	6	Q02660	Q02660 bos taurus	768	6	3.5	1165	1	Q9Y8M2	Q9y8m2 aeropyrum p
696	6	3.5	875	12	Q9WQH4	Q9wqh4 human immun	769	6	3.5	1169	5	Q9NIS0	Q9nis0 toxoplasma
697	6	3.5	877	13	Q9PW88	Q9pw88 carassius a	770	6	3.5	1184	5	Q9U3V9	Q9u3v9 drosophila
698	6	3.5	883	12	Q9WII9	Q9wi19 echovirus 3	771	6	3.5	1185	5	Q9V3V8	Q9v3v8 drosophila
699	6	3.5	883	12	Q9WII9	Q9wi19 echovirus 3	772	6	3.5	1194	5	Q9W485	Q9w485 drosophila
700	6	3.5	888	10	Q04206	Q04206 arabidopsis	773	6	3.5	1197	12	Q69101	Q69101 herpes simp
701	6	3.5	891	1	Q93635	Q93635 thermococcu	774	6	3.5	1202	11	P97607	P97607 rattus norv
702	6	3.5	892	4	Q9V6F6	Q9v6f6 homo sapien	775	6	3.5	1203	10	Q9SN55	Q9sn55 arabidopsis
703	6	3.5	896	12	Q9QEW4	Q9qew4 echovirus 3	776	6	3.5	1211	11	Q9R0L3	Q9r0l3 rattus norv
704	6	3.5	899	10	Q82176	Q82176 arabidopsis	777	6	3.5	1212	12	Q66805	Q66805 echovirus 2
705	6	3.5	899	11	Q9WTK5	Q9wtk5 mus musculu	778	6	3.5	1217	12	Q9YPM8	Q9ypm8 echovirus 3
706	6	3.5	906	4	Q92770	Q92770 homo sapien	779	6	3.5	1222	10	Q9SVW6	Q9svw6 arabidopsis
707	6	3.5	906	4	Q92998	Q92998 homo sapien	780	6	3.5	1235	11	Q9R0L5	Q9r0l5 rattus norv
708	6	3.5	909	2	Q9RT44	Q9rt44 deinococcus	781	6	3.5	1235	12	Q9IBP3	Q9ibp3 western equ
709	6	3.5	911	6	Q9N1F0	Q9n1f0 bos taurus	782	6	3.5	1236	12	Q9JIK1	Q9jik1 western equ
710	6	3.5	915	4	Q9S769	Q9s769 homo sapien	783	6	3.5	1238	12	Q72768	Q72768 eastern equ
711	6	3.5	915	4	Q9P1S8	Q9p1s8 homo sapien	784	6	3.5	1241	12	Q66579	Q66579 eastern equ
712	6	3.5	915	5	Q9XX53	Q9xx53 caenorhabdi	785	6	3.5	1242	12	Q88678	Q88678 eastern equ
713	6	3.5	917	4	Q9U1T2	Q9uit2 homo sapien	786	6	3.5	1242	12	Q88790	Q88790 eastern equ
714	6	3.5	917	4	Q95937	Q95937 homo sapien	787	6	3.5	1242	12	Q88792	Q88792 eastern equ
715	6	3.5	918	11	Q9WVL6	Q9wvl6 mus musculu	788	6	3.5	1242	12	Q88793	Q88793 eastern equ
716	6	3.5	919	11	Q9JMA8	Q9jma8 rattus norv	789	6	3.5	1242	12	Q88794	Q88794 eastern equ
717	6	3.5	920	11	Q35674	Q35674 mus musculu	790	6	3.5	1242	12	Q88795	Q88795 eastern equ
718	6	3.5	920	12	Q88496	Q88496 theiller's e	791	6	3.5	1242	12	Q88796	Q88796 eastern equ
719	6	3.5	920	12	Q88497	Q88497 theiller's e	792	6	3.5	1242	12	Q88797	Q88797 eastern equ
720	6	3.5	921	11	Q9R0D0	Q9r0d0 mus musculu	793	6	3.5	1242	12	Q88798	Q88798 eastern equ
721	6	3.5	922	12	Q56165	Q56165 theiller's e	794	6	3.5	1242	12	Q88799	Q88799 eastern equ
722	6	3.5	922	12	Q88498	Q88498 theiller's e	795	6	3.5	1242	12	Q08359	Q08359 eastern equ
723	6	3.5	923	11	Q9Z0Q2	Q9z0q2 mus musculu	796	6	3.5	1242	12	Q9PZK7	Q9pzx7 eastern equ
724	6	3.5	925	10	Q9SIX4	Q9s1x4 arabidopsis	797	6	3.5	1242	12	Q9PZX6	Q9pzx6 eastern equ
725	6	3.5	930	6	Q9XSW5	Q9xsw5 bos taurus	798	6	3.5	1242	12	Q9PZX5	Q9pzx5 eastern equ
726	6	3.5	932	11	Q35802	Q35802 rattus norv	799	6	3.5	1242	12	Q9PZX4	Q9pzx4 eastern equ
727	6	3.5	934	2	Q9KDU7	Q9kdu7 bacillus ha	800	6	3.5	1242	12	Q9PZX3	Q9pzx3 eastern equ
728	6	3.5	936	5	Q9N910	Q9n910 trypanosoma	801	6	3.5	1242	12	Q9PZX2	Q9pzx2 eastern equ
729	6	3.5	937	12	Q40949	Q40949 kaposi's sa	802	6	3.5	1242	12	Q9PZX1	Q9pzx1 eastern equ
730	6	3.5	942	2	P74081	P74081 synchocyst	803	6	3.5	1242	12	Q9PZX0	Q9pzx0 eastern equ
731	6	3.5	945	4	Q9P2L6	Q9p2l6 homo sapien	804	6	3.5	1242	12	Q9PZW9	Q9pzw9 eastern equ
732	6	3.5	946	3	Q13284	Q13284 phaeosphaer	805	6	3.5	1242	12	Q9PZW8	Q9pzw8 eastern equ
733	6	3.5	946	4	Q9UGE8	Q9uge8 homo sapien	806	6	3.5	1242	12	Q9PZW7	Q9pzw7 eastern equ
734	6	3.5	955	4	Q99466	Q99466 homo sapien	807	6	3.5	1242	12	Q9PZW6	Q9pzw6 eastern equ
735	6	3.5	955	11	Q9QX37	Q9qy37 mus musculu	808	6	3.5	1247	11	Q9QVE5	Q9qve5 mus musculu
736	6	3.5	962	5	Q9VRV1	Q9vr1 drosophila	809	6	3.5	1258	4	Q9P2A7	Q9p2a7 homo sapien
737	6	3.5	967	2	Q9KZD5	Q9kzd5 streptomyce	810	6	3.5	1273	11	Q9R0L4	Q9r0l4 rattus norv
738	6	3.5	971	5	Q45785	Q45785 caenorhabdi	811	6	3.5	1287	5	Q62245	Q62245 caenorhabdi
739	6	3.5	975	5	Q9V7L9	Q9v7l9 drosophila	812	6	3.5	1288	3	O13682	O13682 schizosacch
740	6	3.5	977	1	Q9Y9V1	Q9y9v1 aeropyrum p	813	6	3.5	1294	10	Q9S814	Q9s814 arabidopsis
741	6	3.5	983	10	Q9L134	Q9l134 oryza sativ	814	6	3.5	1296	12	P88964	P88964 kaposi's sa
742	6	3.5	987	2	Q9XBE0	Q9xbe0 amycolatops	815	6	3.5	1296	12	Q9QR70	Q9qr70 kaposi's sa
743	6	3.5	987	5	Q9WD98	Q9wd98 drosophila	816	6	3.5	1297	4	O15025	O15025 homo sapien
744	6	3.5	1010	4	Q9UDF0	Q9ujf0 homo sapien	817	6	3.5	1300	11	Q9WTL4	Q9wtl4 mus musculu
745	6	3.5	1011	2	P74176	P74176 synchocyst	818	6	3.5	1303	12	P88967	P88967 kaposi's sa
746	6	3.5	1015	5	Q9NNW2	Q9nmw2 leishmania	819	6	3.5	1308	12	Q98145	Q98145 kaposi's sa
747	6	3.5	1016	2	O59206	O59206 bradyrhizob	820	6	3.5	1309	2	Q9KS77	Q9ks77 vibrio chol
748	6	3.5	1018	5	O17874	O17874 caenorhabdi	821	6	3.5	1315	10	Q82635	Q82635 arabidopsis
749	6	3.5	1024	5	Q99570	Q99570 caenorhabdi	822	6	3.5	1323	10	Q9M0M2	Q9m0m2 arabidopsis

823	6	3.5	1344	12	Q88483
824	6	3.5	1360	10	Q9STM3
825	6	3.5	1410	2	Q9K1Z9
826	6	3.5	1416	12	Q92812
827	6	3.5	1451	5	Q01737
828	6	3.5	1453	10	Q9LQ02
829	6	3.5	1455	5	Q9V587
830	6	3.5	1469	2	Q9X6V1
831	6	3.5	1513	11	Q62635
832	6	3.5	1551	5	Q9NGV4
833	6	3.5	1556	10	Q9MB97
834	6	3.5	1559	10	Q49889
835	6	3.5	1644	12	Q9NR76
836	6	3.5	1718	5	Q9NBE2
837	6	3.5	1756	5	Q9VBE4
838	6	3.5	1789	12	Q83883
839	6	3.5	1839	2	Q30765
840	6	3.5	1911	5	Q9W4M7
841	6	3.5	1977	4	Q15858
842	6	3.5	1984	6	Q28644
843	6	3.5	1984	11	Q08562
844	6	3.5	1999	4	Q99940
845	6	3.5	2003	4	Q00306
846	6	3.5	2033	4	Q92817
847	6	3.5	2082	2	Q9S200
848	6	3.5	2187	12	Q9YL65
849	6	3.5	2187	12	Q9QSN9
850	6	3.5	2194	12	Q9WN78
851	6	3.5	2224	4	Q43737
852	6	3.5	2261	4	Q9NOV4
853	6	3.5	2286	4	Q9Y5S4
854	6	3.5	2297	2	Q54155
855	6	3.5	2444	5	Q9VP88
856	6	3.5	2478	13	Q93406
857	6	3.5	2487	6	Q9N1T0
858	6	3.5	2602	4	Q9UEV9
859	6	3.5	2602	4	Q75369
860	6	3.5	2618	5	Q9N6P9
861	6	3.5	2635	12	P88955
862	6	3.5	2635	12	Q40942
863	6	3.5	2703	10	Q9MAZ4
864	6	3.5	3013	12	Q92530
865	6	3.5	3016	12	Q92531
866	6	3.5	3033	12	Q91Z66
867	6	3.5	3033	12	Q91Z66
868	6	3.5	3232	3	Q94205
869	6	3.5	3396	5	Q9W555
870	6	3.5	4861	4	Q15751
871	6	3.5	5376	11	Q88799
872	6	3.5	5590	2	Q9RLP6

ALIGNMENTS

[illegible]

```
RC STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Bueli C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Cressy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eissen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome II of Arabidopsis thaliana.";
RL Nature 402:761-768(1999);
DR EMBL; AC007661; AAD32766.1; -.
SQ SEQUENCE 561 AA; 61232 MW; EEFB0BF3127E7680 CRC64;

Query Match 5.3%; Score 9; DB 10; Length 561;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 GPCLSLCL 14
DB 405 GPCLSLCL 413
|||||

RESULT 3
Q9JP96 PRELIMINARY; PRT; 440 AA.
ID Q9JP96
AC Q9JP96
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE ORF440 PROTEIN.
OS Rhodocyclus gelatinosus (Rhodopseudomonas gelatinosa).
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae;
OC Rubrivivax.
OX NCBI_TaxID=28068;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TL144;
RA Nagashima K.V., Shimada K., Matsuura K.;
RT "Phylogenetic analysis of photosynthetic genes of Rhodocyclus
RT gelatinosus: Possibility of horizontal gene transfer in purple
RT bacteria.";
RL Photosyn. Res. 36:185-191(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TL144;
RX MEDLINE=94132007; PubMed=8300574;
RA Nagashima K.V., Matsuura K., Ohyama S., Shimada K.;
RT "Primary structure and transcription of genes encoding B870 and
RT photosynthetic reaction center apoproteins from Rubrivivax
RT gelatinosus.";
RL J. Biol. Chem. 269:2477-2484(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=TL144;
RA Igarashi N., Shimada K., Matsuura K., Nagashima K.V.;
RT "PHOTOSYNTHETIC GENE CLUSTER IN PURPLE BACTERIUM, RUBRIVIVAX
RT GELATINOSUS.";
RL (in) Garab G. (eds.);
RL Photosynthesis. Mechanisms and Effects IV:2889-2892;
RL Kluwer Academic Publishing (1998).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=TL144;
RA Menin L., Yoshida M., Jaquinod M., Nagashima K.V., Matsuura K.,
RA Farot P., Vermeglio A.;
RT "Dark aerobic growth conditions induce the synthesis of a high
RT midpoint potential cytochrome c8 in the photosynthetic bacterium
RT Rubrivivax gelatinosus.";
RL Biochemistry 0:0-0(1999).
DR EMBL; AB034704; BAA94065.1; -.
SQ SEQUENCE 440 AA; 45186 MW; 53238733BAF92A34 CRC64;
```

```
Query Match 4.7%; Score 8; DB 2; Length 440;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 LALPALLP 23
DB 42 LALPALLP 49
|||||

RESULT 4
O76201 PRELIMINARY; PRT; 82 AA.
ID O76201
AC O76201;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE NEUROTOXIN TX3-2 PRECURSOR.
GN TX3-2.
OS Phoneutria nigriventer (Brazilian armed spider).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC Araneomorphae; Entelegynae; Lycosoidea; Ctenidae; Phoneutria.
OX NCBI_TaxID=6918;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=VENOM GLAND;
RX MEDLINE=99053416; PubMed=9839681;
RA Kalapothakis E., Penaforte C.L., Leao R.M., Cruz J.S., Prado V.F.,
RA Cordeiro M.N., Diniz C.R., Romano-Silva M.A., Prado M.A.M.,
RA Gomez M.V., Beirao P.S.L.;
RT "Cloning, cDNA sequence analysis and patch clamp studies of a toxin
RT from the venom of the armed spider (Phoneutria nigriventer).";
RL Toxicon 36:1971-1980(1998).
RN [2]
RP SEQUENCE OF 38-71.
RC TISSUE=VENOM;
RX MEDLINE=93190315; PubMed=8446961;
RA Cordeiro M. do N., De Figueiredo S.G., Valentim A. do C., Diniz C.R.,
RA von Eickstedt V.R.D., Gilroy J., Richardson M.;
RT "Purification and amino acid sequences of six Tx3 type neurotoxins
RT from the venom of the Brazilian 'armed' spider Phoneutria nigriventer
RT (Keys).";
RL Toxicon 31:35-42(1993).
CC -!- FUNCTION: ANTAGONIST OF L-TYPE CALCIUM CHANNELS. INDUCES IMMEDIATE
CC CLOCKWISE GYRATION AND FLACCID PARALYSIS AFTER 6 HOURS AT DOSE
CC LEVELS OF 5 MICROGRAMS PER MOUSE.
DR EMBL; AF015664; AAC26168.1; -.
KW Venom; Neurotoxin; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT PROPEP 22 37
FT CHAIN 38 71 NEUROTOXIN TX3-2.
FT PROPEP 72 82
SQ SEQUENCE 82 AA; 8927 MW; OCA4E00F5AAAD8B9 CRC64;

Query Match 4.1%; Score 7; DB 5; Length 82;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 99 EEEARAC 105
DB 33 EEEARAC 39
|||||

RESULT 5
O69290 PRELIMINARY; PRT; 83 AA.
ID O69290
AC O69290;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE GALLID HERPESVIRUS ANTISENSE RNA.
OS Gallid herpesvirus 1.
```


OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=10386;

EN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=94292920; PubMed=8021600;

Li D.S., Pastorek J., Zelnik V., Smith G.D., Ross L.J.;

"Identification of novel transcripts complementary to the Marek's
 RT disease virus homologue of the ICP4 gene of herpes simplex virus.";

RL J. Gen. Virol. 75:1713-1722(1994).

DR EMBL; L29643; AAA62479.1; -

SQ SEQUENCE 83 AA; 8825 MW; C8B01AA0A4FB8D5C CRC64;

Query Match 4.1%; Score 7; DB 12; Length 83;

Best Local Similarity 100.0%; Pred. No. 21;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 151 PRSGRR 157

Db 75 PRSGRR 81

RESULT 6

Q9PEH3

AC Q9PEH3; PRELIMINARY; PRT; 106 AA.

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)

DE HYPOTHETICAL PROTEIN XF1055.

GN XF1055.

OS Xylella fastidiosa.

OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;

OC Xylella.

OX NCBI_TaxID=2371;

EN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=9A5C;

RA MEDLINE=20365717; PubMed=10910347;

RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,

Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,

Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,

Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,

Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,

Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,

Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,

Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,

Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,

Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,

Krieger J.E., Kuranae E.E., Laigret F., Lambais M.R., Leite L.C.C.,

Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,

Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,

Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,

Menck C.F.M., Miracca E.C., Miyaki C.Y., Montello-Vitorello C.B.,

Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,

Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,

de Oliveira M.C., de Oliveira R.C., Palmeri D.A., Paris A.,

Peixoto B.R., Pereira A.G., Pereira H.A. Jr., Pesquero J.B.,

Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,

de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,

da Silva A.C.R., da Silva F.R., Silva W.A. Jr.,

da Silva J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,

de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshuko M.H.,

Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,

Zago M.A., Zatz M., Meidanis J., Setubal J.C.;

"The genome sequence of the plant pathogen Xylella fastidiosa.";

RL Nature 406:151-157(2000).

DR EMBL; AE003942; AAF83865.1; -

KW Hypothetical protein.

SQ SEQUENCE 106 AA; 12175 MW; 9E048C950D8A6959 CRC64;

Query Match 4.1%; Score 7; DB 2;

Best Local Similarity 100.0%; Pred. No. 26;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 RDAETGE 45

Db 87 RDAETGE 93

RESULT 7

Q29173

ID Q29173; PRELIMINARY; PRT; 115 AA.

AC Q29173;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-OCT-2000 (TrEMBLrel. 13, Last annotation update)

DE CYTOPLASMIC BETA-ACTIN (FRAGMENT).

OS Sus scrofa (pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OX NCBI_TaxID=9823;

EN [1]

RP SEQUENCE FROM N.A.

RX TISSUE=SMALL INTESTINE;

RA MEDLINE=96327607; PubMed=8672129;

Winteroe A.K., Fredholm M., Davies W.;

"Evaluation and characterization of a porcine small intestine cDNA

library: analysis of 839 clones.";

RL Mamm. Genome 7:509-517(1996).

DR EMBL; F14786; CAA23255.1; -

DR INTERPRO; IPR000279; -

DR PFAM; PF00022; actin; 1.

FT NON_TER 1

FT NON_TER 115

SO SEQUENCE 115 AA; 13566 MW; 1891FC8A9FBA149B CRC64;

Query Match 4.1%; Score 7; DB 6;

Best Local Similarity 100.0%; Pred. No. 28;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 ALPALLP 23

Db 56 ALPALLP 62

RESULT 8

Q9UVF0

ID Q9UVF0; PRELIMINARY; PRT; 123 AA.

AC Q9UVF0;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)

DE HYPOTHETICAL 13.6 KDA PROTEIN.

OS 51W.

OC Yarrowia lipolytica (Candida lipolytica).

OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;

OC Dipodascaceae; Yarrowia.

OX NCBI_TaxID=4952;

EN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 20460 (W29); TISSUE=MYCELIUM;

Sanchez M., Dominguez A.;

"Gene order in a 10275 bp fragment from Yarrowia lipolytica including

YlURA5 and YlISEG65 adjacent genes conserved in four yeast species.";

RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ006754; CAB55336.1; -

KW Hypothetical protein.

SQ SEQUENCE 123 AA; 13566 MW; 5ED1BBF5FC79E543 CRC64;

Query Match 4.1%; Score 7; DB 3;

Best Local Similarity 100.0%; Pred. No. 29;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LCLVIAL 18
|||||
DB 66 LCLVIAL 72

RESULT 9
Q9VEK1 PRELIMINARY; PRT; 124 AA.
AC Q9VEK1
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE CG14332 PROTEIN.
GN CG14332.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Folsler C., Gabrielson A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwa C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Mosnere A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
DR EMBL; AE003717; AAF55420.1;
DR FLYBASE; FBgn0038509; CG14332.
SQ SEQUENCE 124 AA; 12630 MW; A0092E5D3E2EC585 CRC64;

Query Match 4.1%; Score 7; DB 5; Length 124;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 147 RGGAPRS 153

DB 88 RGGAPRS 94
|||||

RESULT 10
Q9RJG5 PRELIMINARY; PRT; 173 AA.
AC Q9RJG5
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE HYPOTHETICAL 18.2 KDA PROTEIN.
GN SCF76.09.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Murphy L., Harris D.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL121600; CAB56729.1;
KW Hypothetical protein.
SQ SEQUENCE 173 AA; 18218 MW; 2D67DB6EA42F3CE4 CRC64;

Query Match 4.1%; Score 7; DB 2; Length 173;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 AVRGVAE 32
|||||
DB 43 AVRGVAE 49

RESULT 11
Q9L7T6 PRELIMINARY; PRT; 239 AA.
AC Q9L7T6
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE POLAR FLAGELLAR MOTOR PROTEIN MOPB.
GN MOPB.
OS Rhodospirillum centenum (Rhodocista centenaria).
OC Bacteria; Proteobacteria; alpha subdivision; Rhodospirillaceae.
OX NCBI_TaxID=34018;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC51521;
RA McClain J.R., Rollo D., Bauer C.E., Rushing B.R., Jiang Z.;
RT "Gene duplication and lateral transfer events giving rise to
Rhodospirillum centenum polar and lateral flagellar motor switch
components";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF220007; AAF35853.1;
KW Flagella.
SQ SEQUENCE 239 AA; 25615 MW; ADA9232F477D3662 CRC64;

Query Match 4.1%; Score 7; DB 2; Length 239;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 LVLALPA 20
DB 118 LVLALPA 124

RESULT 12
Q9SF29 PRELIMINARY; PRT; 266 AA.
ID Q9SF29
AC Q9SF29
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)
DE FLIF8.33 PROTEIN.
GN FLIF8.33.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
RA Rowan C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome I BAC FlIF8 genomic sequence.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC016661; AAF23303.1; -;
DR INTERPRO: IPR000727; -;
SQ SEQUENCE 266 AA; 29983 MW; EA30258813FA850C CRC64;

Query Match 4.1%; Score 7; DB 10; Length 266;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 LVLALPA 20
DB 109 LVLALPA 115

RESULT 13
Q04077 PRELIMINARY; PRT; 271 AA.
ID Q04077
AC Q04077
DT 01-JUN-1998 (TRENBLrel. 06, Created)
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE TROPONIN I (FRAGMENT).
OS Chlamys nipponensis akazara (Akazara scallop) (Japanese scallop).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Pectinoidea;
OC Pectinoidea; Pectinidae; Chlamys.
OX NCBI_TaxID=6571;
RN [1]
RP SEQUENCE FROM N.A.
RA Nishita K., Ojima T., Soejima T.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB009368; BAA23775.1; -;
DR INTERPRO: IPR001978; -;
DR PFAM: PF00992; Troponin; 1.
FT NON_TER
SQ SEQUENCE 271 AA; 32068 MW; 83159097F86E44AF CRC64;

Query Match 4.1%; Score 7; DB 5; Length 271;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 EREEAR 103
DB 100 EREEAR 106

RESULT 14
Q9ZDN1 PRELIMINARY; PRT; 274 AA.
ID Q9ZDN1
AC Q9ZDN1
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE HYPOTHETICAL 31.4 KDA PROTEIN.
GN RP296.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MADRID E;
RX MEDLINE=98039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sichertitz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria.";
RL Nature 396:133-140(1998).
DR EMBL: AJ235271; CAA14757.1; -;
KW Hypothetical protein.
SQ SEQUENCE 274 AA; 31365 MW; 02D81B5393C7D9A8 CRC64;

Query Match 4.1%; Score 7; DB 2; Length 274;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 LSLCLIV 15
DB 9 LSLCLIV 15

RESULT 15
Q9LGS9 PRELIMINARY; PRT; 286 AA.
ID Q9LGS9
AC Q9LGS9
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE P0489A01.12 PROTEIN.
GN P0489A01.12.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0489A01.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP002484; BAA9521.1; -;
SQ SEQUENCE 286 AA; 30175 MW; 43E98C837C3D8944 CRC64;

Query Match 4.1%; Score 7; DB 10; Length 286;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 GEREEA 102
DB 131 GEREEA 137

```

Query Match          4.1%; Score 7; DB 5; Length 293;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 97 EREEAR 103
    |||||
b 122 EREEAR 128

RESULT 18
9SM72 PRELIMINARY; PRT; 294 AA.
C Q9SM72;
T 01-MAY-2000 (TrEMBLrel. 13, Created)
T 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
T 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
E ZB0013.1.
S Oryza sativa (Rice).
C Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
C Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.
C NCBI_TaxID=4530;
N [1]
N SEQUENCE FROM N.A.
C STRAIN=CV. INDICA;
C Hong g.; Zhou b.;
T "Oryza sativa genomic DNA chromosome 4, clone: tl7804.";
L Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
R EMBL; AL117265; CAB55421.1; -.
Q SEQUENCE 294 AA; 32257 MW; FE2CF7CBCAD33B2C CRC64;

Query Match          4.1%; Score 7; DB 10; Length 294;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 147 RGGAPRS 153
    |||||
b 3 RGGAPRS 9

RESULT 19
9NNNF4 PRELIMINARY; PRT; 299 AA.
C Q9NNF4;
T 01-OCT-2000 (TrEMBLrel. 15, Created)
T 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
T 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
E POSSIBLE TRANSPORT PROTEIN (FRAGMENT).
S LM15.230.
N Leishmania major.
C Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
C NCBI_TaxID=5664;
N [1]
N SEQUENCE FROM N.A.
C STRAIN=FRIEDLIN;
C Murphy L.; Quail M.; Harris D.; Rajandream M.; Ivens A.; Barrell B.;
X Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
R EMBL; AL160371; CAC00256.1; -.
T NON_TER 1
T NON_TER 299
T NON_TER 299
Q SEQUENCE 299 AA; 33076 MW; 36434F23713827E7 CRC64;

Query Match          4.1%; Score 7; DB 5; Length 299;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

y 22 LPVPAVR 28
    |||||
b 285 LPVPAVR 291

```

```

RESULT 20
O15987
ID O15987 PRELIMINARY; PRT; 314 AA.
AC O15987;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE TROPONIN I.
OS Patinopecten yessoensis (Ezo giant scallop) (Vesso scallop).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Pectinoidea;
OC Pectinoidea; Pectinidae; Patinopecten.
OX NCBI_TaxID=6573;
RN [1]
RP SEQUENCE FROM N.A.
RA Hotta A.L.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB008005; BAA22852.1; -.
DR INTERPRO; IPR001978; -.
DR PFAM; PF00992; Troponin; 1.
SQ SEQUENCE 314 AA; 37009 MW; 2FB61D644DBCAED8 CRC64;

Query Match 4.1%; Score 7; DB 5; Length 314;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 97 EREEAR 103
Db 139 EREEAR 145

RESULT 21
O45987
ID O45987 PRELIMINARY; PRT; 329 AA.
AC O45987;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TREMBLrel. 01, Last annotation update)
DE ORF1 (FLAR), (FLAZ), AND ORF2 GENES, COMPLETE CDS (FLAZ).
GN FLAR.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=76;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CB15;
RA Ely B.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U27302; AAC62520.1; -.
SQ SEQUENCE 329 AA; 33796 MW; AEB8C1A06ACC3D37 CRC64;

Query Match 4.1%; Score 7; DB 2; Length 329;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLALPAL 21
Db 245 VLALPAL 251

RESULT 22
Q9PJX7
ID Q9PJX7 PRELIMINARY; PRT; 335 AA.
AC Q9PJX7;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE GTP-BINDING PROTEIN, GTP1/OBG FAMILY.
GN TC0699.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.

```

```

OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MOPN / NIGG;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Read T.D., Brunham R., Shen C., Gill S.R., Heidelberg J.F., White O.,
RA Hickey E.K., Peterson J., Umayam L.A., Utterback T., Berry K.,
RA Bass S., Linher K., Weidman J., Khouri H., Craven B., Bowman C.,
RA Dodson R., Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G.,
RA Salzberg S.L., Eisen J., Fraser C.M.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE002339; AAF73594.1; -.
DR TIGR; TC0699; -.
DR INTERPRO; IPR000765; -.
DR PFAM; PF01018; GTP1_OBG; 1.
DR PRINTS; PR00326; GTP1OBG.
SQ SEQUENCE 335 AA; 36766 MW; 987057B5C95D9466 CRC64;

Query Match 4.1%; Score 7; DB 2; Length 335;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 RDAETGE 45
Db 96 RDAETGE 102

RESULT 23
Q9Z808
ID Q9Z808 PRELIMINARY; PRT; 335 AA.
AC Q9Z808;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE GTP BINDING PROTEIN.
GN YHBZ.
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=9920606; PubMed=10192388;
RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grinwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
DR EMBL; AE001639; AAD18684.1; -.
DR EMBL; AP002547; BAA98750.1; -.
DR INTERPRO; IPR000765; -.
DR INTERPRO; IPR000834; -.
DR PFAM; PF01018; GTP1_OBG; 1.
DR PRINTS; PR00326; GTP1OBG.

```

```
DR PROSITE; PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_1.
SQ SEQUENCE 335 AA; 36727 MW; 322223517B94E871 CRC64;

Query Match
Best Local Similarity 4.1%; Score 7; DB 2; Length 335;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 RDAETGE 45
DB 96 RDAETGE 102
|||||

RESULT 24
Q92954 PRELIMINARY; PRT; 338 AA.
AC Q92954;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE CATIONIC AMINO ACID TRANSPORTER.
GN YTFE OR CP0645.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CWL029;
RX MEDLINE=99206606; PubMed=10192388;
RA Kaiman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis."
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=J138;
RX MEDLINE=20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA".
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brucham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Unayam L.A., Utterback T.,
RA Berry K., Bass S., Linher K., Weidman J., Khouri H., Craven B.,
RA Bowman C., Dodson R., Gwinn M., Nelson W., DeBoy R., Kolonay J.,
RA McClarty G., Salzberg S.L., Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406(2000).
DR EMBL; AP002545; BAA98338.1; -
DR EMBL; AE002222; AAF38460.1; -
DR TIGR; CP0645; -
SQ SEQUENCE 338 AA; 36939 MW; AA4A69E12A20CB9C CRC64;

Query Match
Best Local Similarity 4.1%; Score 7; DB 2; Length 338;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 SLCLVL 16
DB 313 SLCLVL 319
|||||

RESULT 25
Q9NAL9 PRELIMINARY; PRT; 338 AA.
ID Q9NAL9
```

```
AC Q9NAL9;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE Y105C5B.23 PROTEIN.
GN Y105C5B.23.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurray A.A.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; AL110479; CAB54371.1; -
SQ SEQUENCE 338 AA; 39312 MW; A5EF86EED3B6B3CD CRC64;

Query Match
Best Local Similarity 4.1%; Score 7; DB 5; Length 338;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 SLCLVL 16
DB 30 SLCLVL 36
|||||

RESULT 26
Q9K2C3 PRELIMINARY; PRT; 343 AA.
ID Q9K2C3
AC Q9K2C3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE GTP1/OBG FAMILY PROTEIN.
GN CP0208.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AR39;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brucham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Unayam L.A., Utterback T.,
RA Berry K., Bass S., Linher K., Weidman J., Khouri H., Craven B.,
RA Bowman C., Dodson R., Gwinn M., Nelson W., DeBoy R., Kolonay J.,
RA McClarty G., Salzberg S.L., Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406(2000).
DR EMBL; AE002181; AAF38079.1; -
DR TIGR; CP0208; -
SQ SEQUENCE 343 AA; 37684 MW; 504146CEA5F6D297 CRC64;

Query Match
Best Local Similarity 4.1%; Score 7; DB 2; Length 343;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 RDAETGE 45
DB 104 RDAETGE 110
|||||

RESULT 27
Q9JL20
```

```

ID Q9JL20 PRELIMINARY; PRT; 360 AA.
AC Q9JL20;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE CC CHEMOKINE RECEPTOR 10B (FRAGMENT).
GN CCR10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=20191997; PubMed=10725696;
RA Jarmin D.I., Rits M., Bota D., Gerard N.P., Graham G.J.,
RA Clark-Lewis I., Gerard C.;
RT "Cutting edge: identification of the orphan receptor G-protein-coupled
RT receptor 2 as CCR10, a specific receptor for the chemokine ESKine.";
RL J. Immunol. 164:3460-3464(2000).
DR EMBL: AF215983; AAF63711.1; -.
KW Receptor.
FT NON_TER 1
SQ SEQUENCE 360 AA; 38738 MW; 16CBFEF69F785BA23 CRC64;

Query Match 4.1%; Score 7; DB 11; Length 360;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LALPALL 22
DB 170 LALPALL 176

RESULT 28
Q9NZG2 PRELIMINARY; PRT; 361 AA.
ID Q9NZG2
AC Q9NZG2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE 7-TRANSMEMBRANE G-PROTEIN COUPLED RECEPTOR 2.
GN GPR2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20191998; PubMed=10725697;
RA Honey B., Wang W., Buchanan M., Wisenborn A., Soto H., Catron D.,
RA Orozco R., Ruzicka T., Lehmann P., Oldham E., Zlotnik A.;
RT "Cutting edge: the orphan chemokine receptor G protein-coupled
RT receptor-2 (GPR-2, CCR10) binds the skin-associated chemokine CCL27
RT (CTACK/ALP/ILC).";
RL J. Immunol. 164:3465-3470(2000).
DR EMBL: AF208237; AAF72871.1; -.
KW Receptor; Transmembrane.
SQ SEQUENCE 361 AA; 38302 MW; F8AA749118E3926F CRC64;

Query Match 4.1%; Score 7; DB 4; Length 361;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LALPALL 22
DB 172 LALPALL 178

RESULT 29
Q9JL21 PRELIMINARY; PRT; 362 AA.
ID Q9JL21
AC Q9JL21;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE CC CHEMOKINE RECEPTOR 10A.
GN CCR10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=20191997; PubMed=10725696;
RA Jarmin D.I., Rits M., Bota D., Gerard N.P., Graham G.J.,
RA Clark-Lewis I., Gerard C.;
RT "Cutting edge: identification of the orphan receptor G-protein-coupled
RT receptor 2 as CCR10, a specific receptor for the chemokine ESKine.";
RL J. Immunol. 164:3460-3464(2000).
DR EMBL: AF215982; AAF63710.1; -.
KW Receptor.
SQ SEQUENCE 362 AA; 38926 MW; 0FF4EA2005B94E99 CRC64;

Query Match 4.1%; Score 7; DB 11; Length 362;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LALPALL 22
DB 172 LALPALL 178

RESULT 30
Q9JIP1 PRELIMINARY; PRT; 362 AA.
ID Q9JIP1
AC Q9JIP1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE 7-TRANSMEMBRANE G-PROTEIN COUPLED RECEPTOR 2.
GN GPR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang W., Soto H., Oldham E.R., Buchanan M.E., Honey B., Catron D.,
RA Jenkins N., Copeland N.G., Gilbert D.J., Nguyen N., Abrams J.,
RA Kerstenovich D., Smith K., McLanahan T., Vicari A.P., Zlotnik A.;
RT "Identification of a novel chemokine (CCL28), which binds CCR10
RT (GPR2).";
RL J. Biol. Chem. 275:22313-22323(2000).
DR EMBL: AF208238; AAF72872.1; -.
KW Receptor; Transmembrane.
SQ SEQUENCE 362 AA; 38899 MW; 13C4292859C376CE CRC64;

Query Match 4.1%; Score 7; DB 11; Length 362;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LALPALL 22
DB 172 LALPALL 178

RESULT 31
Q94124 PRELIMINARY; PRT; 364 AA.
ID Q94124
AC Q94124;

```

```

DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE STRESS PROTEIN.
OS Rhizopus stolonifer (Rhizopus nigricans).
OC Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
OC Rhizopus.
OC NCBI_TaxID=4846;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 6227B;
RX MEDLINE=99003515; PubMed=9784403;
RA Cresnar B., Plaper A., Breskvar K., Hudnik-Plevnik T.;
RT "cDNA sequence and deduced amino acid sequence of a fungal stress
protein induced in Rhizopus nigricans by steroids.";
RL Biochem. Biophys. Res. Commun. 250:664-667(1998).
DR EMBL; Y10414; CAA71440.1; -.
DR INTERPRO; IPR001823; -.
DR PROSITE; PS00545; ALDOSE_1_EPIMERASE; 1.
DR PRODOM; PD006407; -.
SQ SEQUENCE 364 AA; 40275 MW; 6A2D52DE54B6B56E CRC64;

Query Match 4.1%; Score 7; DB 3; Length 364;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 HAGFCLE 128
DB 320 HAGFCLE 326
|||||

RESULT 32
Q9RJF6 PRELIMINARY; PRT; 374 AA.
ID Q9RJF6
AC Q9RJF6;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE PUTATIVE MEMBRANE PROTEIN.
GN SCF76.18C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
OC NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Murphy L., Harris D.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Redenbach M., Kieser H.M., Denapalae D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL121600; CAB56738.1; -.
SQ SEQUENCE 374 AA; 39594 MW; CE3DC623D4E9DD0C CRC64;

Query Match 4.1%; Score 7; DB 2; Length 374;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALEGP 7
|||||
```

```

Db 343 MRALEGP 349
|||||

RESULT 33
Q9YBB6 PRELIMINARY; PRT; 393 AA.
ID Q9YBB6
AC Q9YBB6;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE 393AA LONG HYPOTHETICAL ACYL-COA DEHYDROGENASE.
GN APE1681.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;
OC Aeropyrum.
OC NCBI_TaxID=56636;
OX [1]
RN SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Negai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushiida N., Oguchi A., Aoki K., Kubota K., Nakamura Y.,
RA Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
Crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000062; BAA80682.1; -.
DR HSSP; Q06319; 1BUC.
DR INTERPRO; IPR001552; -.
DR PFAM; PF00441; Acyl-CoA_dh; 1.
DR PROSITE; PS00073; ACYL_COA_DH_2; UNKNOWN_1.
SQ SEQUENCE 393 AA; 43027 MW; 541D141EF2E8DB8 CRC64;

Query Match 4.1%; Score 7; DB 1; Length 393;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 GPGLSLL 12
DB 71 GPGLSLL 77
|||||

RESULT 34
Q9ZCV6 PRELIMINARY; PRT; 407 AA.
ID Q9ZCV6
AC Q9ZCV6;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
DE BICYCLOMYCIN RESISTANCE PROTEIN (BCRI).
GN RP603.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OC NCBI_TaxID=782;
OX [1]
RN SEQUENCE FROM N.A.
RC STRAIN=MADRID E;
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sacheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
mitochondria.";
RL Nature 396:133-140(1998).
DR EMBL; AJ235272; CAA15047.1; -.
DR INTERPRO; IPR001066; -.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
```


SQ SEQUENCE 407 AA; 45749 MW; 8A5EC5A0BEEF6EC6 CRC64;

Query Match 4.1%; Score 7; DB 2; Length 407;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 SLCLVL 16
| | | | |
DB 378 SLCLVL 384

RESULT 35

Q9M231 ID Q9M231 PRELIMINARY; PRT; 414 AA.
AC Q9M231;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE HYPOTHETICAL 45.7 KDA PROTEIN.
GN F26B15_30.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Nyakatura G., Fartmann B., Dauner D., Sterr W., Holland R.,
RA Weichselgartner M., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.,
RA Quetier F., Salanoubat M.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; ALI38645; CAB87232.1; -;
KW Hypothetical protein.
SQ SEQUENCE 414 AA; 45727 MW; 748441280707BECB CRC64;

Query Match 4.1%; Score 7; DB 10; Length 414;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 GEREEA 102
| | | | |
DB 56 GEREEA 62

RESULT 36

Q9KX36 ID Q9KX36 PRELIMINARY; PRT; 419 AA.
AC Q9KX36;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE PMOB.
GN PMOB.
OS Methylocystis sp. M.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Methylocystaceae; Methylocystis.
OX NCBI_TaxID=51782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M;
RA Gilbert B., McDonald I.R., Finch R., Stafford G.P., Nielsen A.K.,
RA Murrell J.C.;
RT "Molecular Analysis of the pmo (Particulate Methane Monooxygenase)
RT Operons from Two Type II Methanotrophs.";
RL Appl. Environ. Microbiol. 66:966-975(2000).
DR EMBL; U81596; AAF37897.1; -;
SQ SEQUENCE 419 AA; 45583 MW; 40B108AB418A108B CRC64;

Query Match 4.1%; Score 7; DB 2; Length 419;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 VIAPGES 144
| | | | |
DB 353 VIAPGES 359

RESULT 37

Q9X804 ID Q9X804 PRELIMINARY; PRT; 432 AA.
AC Q9X804;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE HYPOTHETICAL 43.3 KDA PROTEIN.
GN SC6G10.18.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Seeger K., Harris D.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapante D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL049497; CAB39872.1; -;
KW Hypothetical protein.
SQ SEQUENCE 432 AA; 43344 MW; 5467FA467AECLF17 CRC64;

Query Match 4.1%; Score 7; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 NVLCGER 98
| | | | |
DB 204 NVLCGER 210

RESULT 38

Q9X8V6 ID Q9X8V6 PRELIMINARY; PRT; 433 AA.
AC Q9X8V6;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE PUTATIVE SECRETED KINASE.
GN SCH35_50C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Oliver K., Harris D.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RX MEDLINE=97000351; PubMed=8843436;
 RA Regenbach M., Kleser H.M., Denapate D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL; AL078610; CAB44377.1; -.
 KW Kinase.
 SQ SEQUENCE 433 AA; 43776 MW; 81584601BA75EF7C CRC64;

Query Match 4.1%; Score 7; DB 2; Length 433;
 Best Local Similarity 100.0%; Pred. No. 84; Mismatches 0; Indels 0; Gaps 0;
 Matches 7; Conservative 0;

Qy 19 PALLPVP 25
 |||||
 Db 47 PALLPVP 53

RESULT 39

P79825 PRELIMINARY; PRT; 446 AA.
 AC P79825;
 DT 01-MAY-1997 (TREMBlrel. 03, Created)
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE HEMOPEXIN-LIKE PROTEIN (FRAGMENT).
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostei;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE=97142135; PubMed=898369;
 RA Miot S., Duval J., Le Goff P.;
 RT "Molecular cloning of a hemopexin-like cDNA from rainbow trout
 liver.";
 RL DNA Seq. 6:311-318(1996).
 DR EMBL; Z68112; CAA92147.1; -.
 DR HSSP; P20058; 1HXN.
 DR INTERPRO; IPR000585; -.
 DR PFAM; PF00045; hemopexin; 4.
 FT NON_TER 1
 SQ SEQUENCE 446 AA; 50454 MW; 13A8B4EA42B37727 CRC64;

Query Match 4.1%; Score 7; DB 13; Length 446;
 Best Local Similarity 100.0%; Pred. No. 86; Mismatches 0; Indels 0; Gaps 0;
 Matches 7; Conservative 0;

Qy 12 LCLVLAL 18
 |||||
 Db 11 LCLVLAL 17

RESULT 40

Query Match 4.1%; Score 7; DB 5; Length 472;

O66700 PRELIMINARY; PRT; 453 AA.
 AC O66700;
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE HYPOTHETICAL 53.1 KDA PROTEIN.
 GN AQ_378.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificales; Aquificaceae; Aquifex.
 OX NCBI_TaxID=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VF5;
 RX MEDLINE=98196666; PubMed=9537320;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Auway M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 aeolicus.";
 RL Nature 392:353-358(1998).
 DR EMBL; AE000887; AAC06670.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 453 AA; 53052 MW; 41F0BD0FD228FF5F CRC64;

Query Match 4.1%; Score 7; DB 2; Length 453;
 Best Local Similarity 100.0%; Pred. No. 87; Mismatches 0; Indels 0; Gaps 0;
 Matches 7; Conservative 0;

Qy 119 FFAHAGF 125
 |||||
 Db 406 FFAHAGF 412

RESULT 41

Q20406 PRELIMINARY; PRT; 472 AA.
 AC Q20406;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-JAN-1999 (TREMBlrel. 09, Last annotation update)
 DE F44F4.6 PROTEIN.
 GN F44F4.6.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Coles L.;
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Noopra A., Saunders D., Showkeen R.,
 RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans.";
 RL Nature 368:32-38(1994).
 DR EMBL; Z37092; CAA85457.1; -.
 SQ SEQUENCE 472 AA; 55412 MW; 36A18D96D9C6245A CRC64;

Best Local Similarity 100.0%; Pred. No. 90;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 PALLPVP 25
DB 317 PALLPVP 323

RESULT 42

Q9R7S6 Q9R7S6 PRELIMINARY; PRT; 481 AA.
AC Q9R7S6;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE OUTER MEMBRANE USHER PROTEIN PMFC PRECURSOR (FRAGMENT).
GN PMFC.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of Escherichia coli K-12 Genome Corresponding
RT to the 12.7-28.0 min Region on the Linkage Map."
RL DNA Res. 3:137-155(1996).
DR EMBL; D90711; BAA3385.1;
DR INTERPRO; IPR00015;
DR PFAM; PF00577; Usher; 1.
DR PROSITE; PS01151; FIMBRIAL_USHER; 1.
FT NON_TER 481
SQ SEQUENCE 481 AA; 53557 MW; DB11E47A9CC6DDC4 CRC64;

Query Match 4.1%; Score 7; DB 2; Length 481;

Best Local Similarity 100.0%; Pred. No. 92;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 GVIAPGE 143
DB 63 GVIAPGE 69

RESULT 43

Q9L2L1 Q9L2L1 PRELIMINARY; PRT; 485 AA.
AC Q9L2L1;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE PUTATIVE MEMBRANE PROTEIN.
GN SCF42.02C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Oliver K., Harris D.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL137165; CAB69664.1;
SQ SEQUENCE 485 AA; 50944 MW; 63053A5803396D10 CRC64;

Query Match 4.1%; Score 7; DB 2; Length 485;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LCLVLAL 18
DB 387 LCLVLAL 393

RESULT 44

Q9R2U8 Q9R2U8 PRELIMINARY; PRT; 501 AA.
AC Q9R2U8;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE TRANSPOSASE.
GN IS469 OR IS469-ORF.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Murphy L., Harris D.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Kinashi H., Yamasaki M.;
RT "Insertion sequence cluster in Streptomyces coelicolor A3(2).";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL133236; CAB61792.1;
DR EMBL; AB032065; BAA84089.1;
SQ SEQUENCE 501 AA; 56154 MW; DABC96397AF415D1 CRC64;

Query Match 4.1%; Score 7; DB 2; Length 501;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 DAETGER 46
DB 169 DAETGER 175

RESULT 45

```

Q9SN03
ID Q9SN03 PRELIMINARY: PRT: 509 AA.
AC Q9SN03;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE HYPOTHETICAL 59.4 KDA PROTEIN.
GN F3A 230.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Barges M., Collado M.C., Navarro P., Terol J., Perez-Alonso M.,
RA Mewes H.W., Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; ALI32978; CAB62123.1; -.
KW Hypothetical protein.
SQ SEQUENCE 509 AA; 59387 MW; DBDCL1188E230AEA4 CRC64;

Query Match 4.1%; Score 7; DB 10; Length 509;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 EEEARAC 105
Db 157 EEEARAC 163
|||||||

RESULT 46
O52237 PRELIMINARY: PRT: 513 AA.
AC O52237;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE SDEK.
GN SDEK.
OS Myxococcus xanthus.
OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
OC Myxococcales; Cystobacterineae; Myxococcaceae; Myxococcus.
OX NCBI_TaxID=34;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DK101;
RA Garza A.G., Pollack J.S., Harris B.Z., Lee A., Keseler I.,
RA Licking E.F., Singer M.H.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF031084; AAB94135.1; -.
DR INTERPRO; IPR000014; -.
DR INTERPRO; IPR000410; -.
DR PFAM; PF00512; signal; 1.
DR PFAM; PF00389; PAS; 1.
SQ SEQUENCE 513 AA; 55718 MW; E5BBA40A7AA4171D CRC64;

Query Match 4.1%; Score 7; DB 2; Length 513;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 EREEAR 103
Db 269 EREEAR 275
|||||||

RESULT 47

```

```

Q9SW67
ID Q9SW67 PRELIMINARY: PRT: 518 AA.
AC Q9SW67;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE CYTOCHROME P450-LIKE PROTEIN.
GN C7A10.50 OR AT4G37310.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Terry N., Vos P., Heljinen L., Mewes H.W., Mayer K.F.X.,
RA Schueller C.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Rose M., Hempel S., Entian K.-D., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z99707; CAB16771.1; -.
DR EMBL; ALI61591; CAB80397.1; -.
DR INTERPRO; IPR001128; -.
DR INTERPRO; IPR002397; -.
DR INTERPRO; IPR002401; -.
DR INTERPRO; IPR002403; -.
DR PFAM; PF00067; P450; 1.
DR PRINTS; PR00359; BP450.
DR PRINTS; PR00385; P450.
DR PRINTS; PR00463; EP450I.
DR PRINTS; PR00465; EP450IV.
DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN.1.
SQ SEQUENCE 518 AA; 58227 MW; EAA686C30B3D6BBC CRC64;

Query Match 4.1%; Score 7; DB 10; Length 518;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 EGPGLSL 11
Db 481 EGPGLSL 487
|||||||

RESULT 48
Q9ZC50 PRELIMINARY: PRT: 546 AA.
ID Q9ZC50;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE SIMILAR TO YFIU PROTEIN.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=6/69;
RA Buchrieser C., Rusniok C., Couve E., Frangeul L., Billault A.,
RA Kunst F., Carniel E., Glaser P.;
RT "DNA sequence of the 102 kbases unstable region of Yersinia pestis.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.

```

DR EMBL; AL031866; CAA21361.1; -
SQ SEQUENCE 546 AA; 57814 MW; 2BF642DAC20C5258 CRC64;

Query Match 4.1%; Score 7; DB 2; Length 546;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 28 RGVATP 34
|||||
Db 448 RGVATP 454

RESULT 49

Q9NMV7 Q9NMV7 PRELIMINARY; PRT; 560 AA.
AC Q9NMV7;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE POSSIBLE IMMEDIATE-EARLY PROTEIN IE180 (FRAGMENT).
GN LM15.365.
OS Leishmania major.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FRIEDLIN;
RA Murphy L., Quail M., Harris D., Rajandream M., Ivens A., Barrell B.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL160371; CAC00453.1; -
FT NON_TER 1
FT NON_TER 560
ET NON_TER
SQ SEQUENCE 560 AA; 59256 MW; 8A784D647B576D94 CRC64;

Query Match 4.1%; Score 7; DB 5; Length 560;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 98 REEERA 104
|||||
Db 495 REEERA 501

RESULT 50

F74569 F74569 PRELIMINARY; PRT; 600 AA.
AC P74569;
DT 01-FEB-1997 (TReMBLrel. 02, Created)
DT 01-FEB-1997 (TReMBLrel. 02, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE ASPARTATE KINASE.
GN LYSC.
OS Synecocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyaajima N., Hirosewa M., Sugiyura M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
Synecocystis sp. strain PCC6803. II. Sequence determination of the
entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
DR EMBL; D90916; BAA18676.1; -
DR INTERPRO; IPR001048; -
DR INTERPRO; IPR001341; -
DR INTERPRO; IPR002106; -

DR INTERPRO; IPR002912; -
DR PFAM; PF00696; aakinasase; 1.
DR PFAM; PF01842; ACT; 4.
DR PROSITE; PS00324; ASPARTOKINASE; 1.
DR PROSITE; PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
SQ SEQUENCE 600 AA; 63533 MW; B30B7101E9FC4D46 CRC64;

Query Match 4.1%; Score 7; DB 2; Length 600;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 PAVRGVA 31
|||||
Db 433 PAVRGVA 439

Search completed: January 30, 2001, 17:11:09
Job time: 288 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2001, 17:07:07 ; Search time 22.54 seconds
(without alignments)

135.435 Million cell updates/sec

Title: US-09-518-931-4
Perfect score: 170
Sequence: 1 MRALEGPGLSLCLVLALPA.....PRSGRRRCRGQVAGPSLAP 170

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 174772 seqs, 17957048 residues

Word size : 6

Total number of hits satisfying chosen parameters: 84

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6.COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/PTCUS.COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	142	83.5	300	2	US-08-794-796-2
2	8	4.7	1172	1	US-08-313-288B-19
3	7	4.1	1611	2	US-08-804-227C-5
4	6	3.5	9	2	US-08-482-651-10
5	6	3.5	70	3	US-09-188-930-131
6	6	3.5	100	1	US-08-473-981A-11
7	6	3.5	100	2	US-08-474-087-11
8	6	3.5	126	3	US-08-513-974B-28
9	6	3.5	135	2	US-08-757-036-3
10	6	3.5	142	2	US-08-164-292B-20
11	6	3.5	142	3	US-08-845-623-20
12	6	3.5	142	3	US-08-815-927-20
13	6	3.5	144	3	US-08-513-974B-366
14	6	3.5	144	3	US-08-513-974B-369
15	6	3.5	150	2	US-08-851-188-1
16	6	3.5	150	2	US-08-851-188-3
17	6	3.5	159	2	US-08-851-188-4
18	6	3.5	180	2	US-08-791-495-7
19	6	3.5	183	5	5168049-4
20	6	3.5	203	3	US-09-106-182-3
21	6	3.5	206	3	US-08-513-974B-27
22	6	3.5	210	2	US-08-791-495-5
23	6	3.5	220	2	US-08-840-683-8
24	6	3.5	223	3	US-08-513-974B-315
25	6	3.5	223	3	US-08-513-974B-364
26	6	3.5	223	3	US-08-513-974B-368
27	6	3.5	283	4	PCT-US96-12374-2
28	6	3.5	287	1	US-07-971-092-2

29	3.5	287	5	5198342-2	Patent No. 5198342
30	3.5	317	2	US-08-555-722-8	Sequence 8, Appli
31	3.5	346	2	US-08-602-359A-34	Sequence 34, Appl
32	3.5	370	3	US-08-513-974B-26	Sequence 26, Appl
33	3.5	370	3	US-08-513-974B-323	Sequence 323, App
34	3.5	427	3	US-08-448-722A-4	Sequence 4, Appli
35	3.5	434	3	US-09-012-072-4	Sequence 4, Appli
36	3.5	437	3	US-09-073-569-2	Sequence 2, Appli
37	3.5	457	5	5268463-7	Patent No. 5268463
38	3.5	480	2	US-08-425-989B-11	Sequence 11, Appl
39	3.5	498	4	PCT-US94-01101-2	Sequence 2, Appli
40	3.5	531	2	US-08-789-078-3	Sequence 3, Appli
41	3.5	531	2	US-08-752-633-3	Sequence 3, Appli
42	3.5	531	4	PCT-US95-04886-3	Sequence 3, Appli
43	3.5	532	1	US-07-618-286-1	Sequence 1, Appli
44	3.5	532	1	US-08-196-003-3	Sequence 3, Appli
45	3.5	532	2	US-08-483-389-118	Sequence 118, App
46	3.5	532	2	US-08-689-870-12	Sequence 12, Appl
47	3.5	532	3	US-08-933-824-3	Sequence 3, Appli
48	3.5	532	5	5284931-2	Patent No. 5284931
49	3.5	554	1	US-08-106-761-2	Sequence 2, Appli
50	3.5	587	1	US-08-398-008A-2	Sequence 2, Appli
51	3.5	587	1	US-07-955-905A-23	Sequence 23, Appl
52	3.5	587	2	US-08-893-333-2	Sequence 2, Appli
53	3.5	589	2	US-08-453-848-13	Sequence 13, Appl
54	3.5	594	2	US-08-785-310A-6	Sequence 6, Appli
55	3.5	751	2	US-08-836-443-3	Sequence 3, Appli
56	3.5	809	4	PCT-US91-01726-3	Sequence 3, Appli
57	3.5	899	1	US-08-365-689-2	Sequence 2, Appli
58	3.5	899	1	US-08-145-138A-2	Sequence 2, Appli
59	3.5	933	1	US-07-747-781-2	Sequence 2, Appli
60	3.5	933	4	PCT-US92-06888-2	Sequence 2, Appli
61	3.5	983	2	US-08-164-292B-26	Sequence 26, Appl
62	3.5	983	3	US-08-845-623-26	Sequence 26, Appl
63	3.5	983	3	US-08-815-927-26	Sequence 26, Appl
64	3.5	1011	3	US-08-836-325-2	Sequence 2, Appli
65	3.5	1248	3	US-08-882-046-6	Sequence 6, Appli
66	3.5	1294	2	US-08-819-288-3	Sequence 3, Appli
67	3.5	1296	3	US-08-728-603-15	Sequence 15, Appl
68	3.5	1321	1	US-08-261-822A-3	Sequence 3, Appli
69	3.5	1321	4	PCT-US95-07744A-3	Sequence 3, Appli
70	3.5	1375	3	US-08-665-259-26	Sequence 26, Appl
71	3.5	1375	3	US-08-762-500-26	Sequence 26, Appl
72	3.5	1410	3	US-09-335-409-3	Sequence 3, Appli
73	3.5	1503	3	US-08-976-255-14	Sequence 14, Appl
74	3.5	1835	3	US-08-836-325-15	Sequence 15, Appl
75	3.5	1969	3	US-08-836-325-16	Sequence 16, Appl
76	3.5	1984	3	US-08-836-325-10	Sequence 10, Appl
77	3.5	1989	3	US-08-836-325-11	Sequence 11, Appl
78	3.5	1989	3	US-08-836-325-12	Sequence 12, Appl
79	3.5	1996	2	US-08-804-227C-9	Sequence 9, Appli
80	3.5	1996	2	US-08-804-198-3	Sequence 3, Appli
81	3.5	2647	2	US-08-583-562B-8	Sequence 8, Appli
82	3.5	2647	2	US-08-779-113-8	Sequence 8, Appli
83	3.5	3033	1	US-07-925-695-8	Sequence 8, Appli
84	3.5	3033	1	US-07-925-695-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1
US-08-794-796-2
; Sequence 2, Application US/08794796
; Patent No. 5885800
; GENERAL INFORMATION:
; APPLICANT: Emery, John
; APPLICANT: Tan, KB
; APPLICANT: Trunch, Alem
; APPLICANT: Young, Peter
; TITLE OF INVENTION: Tumor Necrosis Related Receptor,
; TITLE OF INVENTION: TR4
; NUMBER OF SEQUENCES: 2

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/794,796
FILING DATE: 04-FEB-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Han, William T
REGISTRATION NUMBER: 34,344
REFERENCE/DOCKET NUMBER: GH50000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5219
TELEFAX: 610-270-4026
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 300 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-794-796-2

Query Match 83.5%; Score 142; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 5.5e-126; Indels 0; Gaps 0;
Matches 142; Conservative 0; Mismatches 0;

QY 1 MRALEGGSLSLCLLVLPVAVRGVAETPTYPWRDAETGERLVCAQCPGTFVQR 60
Db 1 MRALEGGSLSLCLLVLPVAVRGVAETPTYPWRDAETGERLVCAQCPGTFVQR 60
QY 61 PCRRDSTTCGCPPRHYTQFWNYLRCRYCNVLCGEREEERACHATHNRACRGTGF 120
Db 61 PCRRDSTTCGCPPRHYTQFWNYLRCRYCNVLCGEREEERACHATHNRACRGTGF 120
QY 121 AHAGFCLHASCPPGAGVIAPG 142
Db 121 AHAGFCLHASCPPGAGVIAPG 142

RESULT 2
US-08-313-288B-19
Sequence 19, Application US/08313288B
Patent No. 5750502
GENERAL INFORMATION:
APPLICANT: Jessell, Thomas M. and Avihu Klar
TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/313,288B
FILING DATE: January 5, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0526
TELEX:
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 1172 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-313-288B-19

Query Match 4.7%; Score 8; DB 1; Length 1172;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 ALEGGPGLS 10
Db 102 ALEGGPGLS 109

RESULT 3
US-08-804-227C-5
Sequence 5, Application US/08804227C
Patent No. 5876991
GENERAL INFORMATION:
APPLICANT: DeHoff, Bradley S.
APPLICANT: Kuhstoss, Stuart A.
APPLICANT: Rostock, Paul R., Jr.
APPLICANT: Sutton, Kimberly L.
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: THOMAS G. PLANT 1501
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII(DOS) Text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,227C
FILING DATE: February 21, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Plant, Thomas, G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1611 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-804-227C-5

```


Query Match 4.1%; Score 7; DB 2; Length 1611;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 98 REEARA 104
| | | | |
Db 1516 REEARA 1522

RESULT 4
US-08-482-651-10
; Sequence 10, Application US/08482651
; Patent No. 5874409
; GENERAL INFORMATION:
; APPLICANT: Victoria, Edward J.
; APPLICANT: Marquis, David M.
; TITLE OF INVENTION: APL IMMUNOREACTIVE PEPTIDES, CONJUGATES
; TITLE OF INVENTION: THEREOF AND METHODS OF TREATMENT FOR APL ANTIBODY-MEDIATED
; TITLE OF INVENTION: PATHOLOGIES
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,651
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Park, Freddie K.
; REGISTRATION NUMBER: 35,636
; REFERENCE/DOCKET NUMBER: 25231-20061.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: xyz (details pg. 16)
; CLONE: 3B6
US-08-482-651-10

Query Match 3.5%; Score 6; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 CLVLAL 18
| | | | |
Db 1 CLVLAL 6

RESULT 5
US-09-188-930-131
; Sequence 131, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.

; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 131
; LENGTH: 70
; TYPE: PRT
; ORGANISM: mouse
US-09-188-930-131

Query Match 3.5%; Score 6; DB 3; Length 70;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 GPGLSL 11
| | | | |
Db 63 GPGLSL 68

RESULT 6
US-08-473-981A-11
; Sequence 11, Application US/08473981A
; Patent No. 5629162
; GENERAL INFORMATION:
; APPLICANT: deFougerolles, Antonin R
; APPLICANT: Springer, Timothy A
; TITLE OF INVENTION: METHODS OF IDENTIFYING AGENTS WHICH MODULATE
; TITLE OF INVENTION: ICAM-3 BINDING TO LFA-1 (AS AMENDED)
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, N.W. SUITE 600
; CITY: WASHINGTON
; STATE: D. C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,981A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MILONIG, ROBERT C
; REGISTRATION NUMBER: 34,395
; REFERENCE/DOCKET NUMBER: 1011.0560004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 100 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-473-981A-11

Query Match 3.5%; Score 6; DB 1; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

Qy      15 VLALPA 20
       |||||
Db      29 VLALPA 34

RESULT 9
US-08-757-036-3
; Sequence 3, Application US/08757036
; Patent No. 5843668
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: HUMAN SQMI PROTEIN HOMOLOG
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/757.036
; FILING DATE: Herewith
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0170 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 135 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 180233
US-08-757-036-3

```

```

Query Match      3.5%; Score 6; DB 2; Length 135;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      152 RSGRR 157
       |||||
Db      111 RSGRR 116

```

```

RESULT 10
US-08-164-292B-20
; Sequence 20, Application US/08164292B
; Patent No. 5820868
; GENERAL INFORMATION:
; APPLICANT: MITTAL, SURESH K.
; APPLICANT: GRAHAM, FRANK L.
; APPLICANT: PREVEC, LUDVIK
; APPLICANT: BABIUK, LORNE A.
; TITLE OF INVENTION: RECOMBINANT PROTEIN PRODUCTION IN BOVINE

```

```

; TITLE OF INVENTION: ADENOVIRUS EXPRESSION VECTOR SYSTEM
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 345 California Street
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104-2675
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/164,292B
; FILING DATE: 09-DEC-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: GRACEY, NANCY J.
; REGISTRATION NUMBER: 28,216
; REFERENCE/DOCKET NUMBER: 29310-20021.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 677-7000
; TELEFAX: (415) 677-7522
; TELEX: 34-0154
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 142 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-164-292B-20

```

```

Query Match      3.5%; Score 6; DB 2; Length 142;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      6 GPGLSL 11
       |||||
Db      12 GPGLSL 17

```

```

RESULT 11
US-08-845-623-20
; Sequence 20, Application US/08845623A
; Patent No. 6001591
; GENERAL INFORMATION:
; APPLICANT: BABIUK, LORNE A.
; APPLICANT: TIKOO, SURESH K.
; APPLICANT: REDDY, POLICE S.
; TITLE OF INVENTION: BOVINE ADENOVIRUS 3 GENOME
; FILE REFERENCE: 293102002120
; CURRENT APPLICATION NUMBER: US/08/845,623A
; CURRENT FILING DATE: 1997-04-25
; EARLIER APPLICATION NUMBER: 08/164,294
; EARLIER FILING DATE: 1993-12-09
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Bovine adenovirus type 3
US-08-845-623-20

```

```

Query Match      3.5%; Score 6; DB 3; Length 142;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      6 GPGLSL 11
       |||||

```

Db 12 GPGSL 17

RESULT 12

US-08-815-927-20

Sequence 20, Application US/08815927

Patent No. 6086890

GENERAL INFORMATION:

APPLICANT: MITTAL, SURESH K.

APPLICANT: GRAHAM, FRANK L.

APPLICANT: PREVIC, LUDVIG

APPLICANT: BABIUK, LORNE A.

TITLE OF INVENTION: RECOMBINANT PROTEIN PRODUCTION IN BOVINE ADENOVIRUS EXPRESSION

TITLE OF INVENTION: VECTOR SYSTEM

FILE REFERENCE: 293102002101

CURRENT APPLICATION NUMBER: US/08/815,927

CURRENT FILING DATE: 1997-03-13

EARLIER APPLICATION NUMBER: 08/164,294

EARLIER FILING DATE: 1993-12-09

NUMBER OF SEQ ID NOS: 34

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 20

LENGTH: 142

TYPE: PRT

ORGANISM: Bovine adenovirus type 3

US-08-815-927-20

Query Match 3.5%; Score 6; DB 3; Length 142;

Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 GPGSL 11

Db 12 GPGSL 17

RESULT 13

US-08-513-974B-366

Sequence 366, Application US/08513974B

Patent No. 6114139

GENERAL INFORMATION:

APPLICANT: Hinuma, Shuji

APPLICANT: Hosoya, Masaki

APPLICANT: Fujii, Ryo

APPLICANT: Ohtaki, Tetsuya

APPLICANT: Fukusumi, Shoji

APPLICANT: Ohgi, Kazuhiro

TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,

TITLE OF INVENTION: PRODUCTION, AND USE THEREOF

NUMBER OF SEQUENCES: 380

CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

STREET: 130 Water Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/513,974B

FILING DATE: 14-SEP-1995

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP95/01599

FILING DATE: 10-AUG-1995

PRIOR APPLICATION DATA: JP 7-093989

APPLICATION NUMBER: JP 7-093989

FILING DATE: 19-AUG-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 7-057186

FILING DATE: 16-MAR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 7-007177

FILING DATE: 20-JAN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-326611

FILING DATE: 28-DEC-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-270017

FILING DATE: 02-NOV-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-236357

FILING DATE: 30-SEP-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-236356

FILING DATE: 30-SEP-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-189274

FILING DATE: 11-AUG-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-189273

FILING DATE: 11-AUG-1945

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-189272

FILING DATE: 11-AUG-1994

ATTORNEY/AGENT INFORMATION:

NAME: Resnick, David S.

REGISTRATION NUMBER: 34,235

REFERENCE/DOCKET NUMBER: 45753

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 366:

SEQUENCE CHARACTERISTICS:

LENGTH: 144 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-513-974B-366

Query Match 3.5%; Score 6; DB 3; Length 144;

Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLALPA 20

Db 38 VLALPA 43

RESULT 14

US-08-513-974B-369

Sequence 369, Application US/08513974B

Patent No. 6114139

GENERAL INFORMATION:

APPLICANT: Hinuma, Shuji

APPLICANT: Hosoya, Masaki

APPLICANT: Fujii, Ryo

APPLICANT: Ohtaki, Tetsuya

APPLICANT: Fukusumi, Shoji

APPLICANT: Ohgi, Kazuhiro

TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,

TITLE OF INVENTION: PRODUCTION, AND USE THEREOF

NUMBER OF SEQUENCES: 380

CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

STREET: 130 Water Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/513.974B
FILING DATE: 14-SEP-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP95/01599
FILING DATE: 10-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-093989
FILING DATE: 19-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-057186
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-007177
FILING DATE: 20-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-326611
FILING DATE: 28-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-270017
FILING DATE: 02-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-236357
FILING DATE: 30-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-236356
FILING DATE: 30-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-189274
FILING DATE: 11-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-189273
FILING DATE: 11-AUG-1945
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-189272
FILING DATE: 11-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S.
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 45753
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 369:
SEQUENCE CHARACTERISTICS:
LENGTH: 144 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-513-974B-369

Query Match 3.5%; Score 6; DB 3; Length 144;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLALPA 20
Db 38 VLALPA 43

RESULT 15
US-08-851-188-1
; Sequence 1, Application US/08851188
; Patent No. 5925542
; GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Shah, Purvi
TITLE OF INVENTION: NOVEL HUMAN PHOSPHODIESTERASE
TITLE OF INVENTION: DELTA SUBUNIT
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851.188
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0284 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 150 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: 2285337
CLONE: 2285337
US-08-851-188-1

Query Match 3.5%; Score 6; DB 2; Length 150;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 RDAETG 44
Db 23 RDAETG 28

RESULT 16
US-08-851-188-3
; Sequence 3, Application US/08851188
; Patent No. 5925542
; GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Shah, Purvi
TITLE OF INVENTION: NOVEL HUMAN PHOSPHODIESTERASE
TITLE OF INVENTION: DELTA SUBUNIT
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,188
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0284 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 150 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1565306
US-08-851-188-3

Query Match 3.5%; Score 6; DB 2; Length 150;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 RDAETG 44
|||||
DB 23 RDAETG 28

RESULT 17
US-08-851-188-4
Sequence 4, Application US/08851188
Patent No. 5925542
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Shah, Purvi
TITLE OF INVENTION: NOVEL HUMAN PHOSPHODIESTERASE
TITLE OF INVENTION: DELTA SUBUNIT
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,188
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0284 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:
LENGTH: 159 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 540267
US-08-851-188-4

Query Match 3.5%; Score 6; DB 2; Length 159;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 RDAETG 44
|||||
DB 32 RDAETG 37

RESULT 18
US-08-791-495-7
Sequence 7, Application US/08791495
Patent No. 5811519
GENERAL INFORMATION:
APPLICANT: Leth, Bernard
APPLICANT: Lucas, Sophie
APPLICANT: De Smet, Charles
APPLICANT: Godelaine, Daniele
APPLICANT: Boon-Falleur, Thierry
TITLE OF INVENTION: LL-1 TUMOR SPECIFIC GENES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/791,495
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Van Amsterdam, John R.
REGISTRATION NUMBER: 40,212
REFERENCE/DOCKET NUMBER: L0461/7005
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 180 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-791-495-7

Query Match 3.5%; Score 6; DB 2; Length 180;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 147 RGGAPR 152
|||||
DB 57 RGGAPR 62

RESULT 19
5168049-4
;PATENT NO. 5168049
; APPLICANT: MEADE, HARRY M.;GARWIN, JEFFREY L.
; TITLE OF INVENTION: PRODUCTION OF STREPTAVIDIN-LIKE
; POLYPEPTIDES
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/185,329
; FILING DATE: 21-APR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 656.873
; FILING DATE: 02-OCT-1984
; SEQ ID NO:4:
; LENGTH: 183
5168049-4

Query Match 3.5%; Score 6; DB 5; Length 183;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 LEGPGL 9
Db 29 LEGPGL 34

RESULT 20
US-09-106-182-3
; Sequence 3, Application US/09106182
; Patent No. 6046035
; GENERAL INFORMATION:
; APPLICANT: Shi, Yanggu
; APPLICANT: Ruben, Steve
; TITLE OF INVENTION: Cardiostrophin-Like Cytokine
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc
; STREET: 9410 Key West Ave
; CITY: Rockville
; STATE: MD
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/106,182
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/051,053
; FILING DATE: 30-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF385
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 203 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-106-182-3

Query Match 3.5%; Score 6; DB 3; Length 203;

Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 17 ALPALL 22
Db 96 ALPALL 101

RESULT 21
US-08-513-974B-27
; Sequence 27, Application US/08513974B
; Patent No. 6114139
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Hosoya, Masaki
; APPLICANT: Fujii, Ryo
; APPLICANT: Ohtaki, Tetsuya
; APPLICANT: Fukusumi, Shoji
; APPLICANT: Ohgi, Kazuhiro
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,
; AND USE THEREOF
; NUMBER OF SEQUENCES: 380
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/513,974B
; FILING DATE: 14-SEP-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP95/01599
; FILING DATE: 10-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-093989
; FILING DATE: 19-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-057186
; FILING DATE: 16-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-007177
; FILING DATE: 20-JAN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-326611
; FILING DATE: 28-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-270017
; FILING DATE: 02-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-236357
; FILING DATE: 30-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-236356
; FILING DATE: 30-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-189274
; FILING DATE: 11-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-189273
; FILING DATE: 11-AUG-1945
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-189272
; FILING DATE: 11-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Resnick, David S.

REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 45753
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 206 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-513-974B-27

Query Match 3.5%; Score 6; DB 3; Length 206;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLALPA 20
|||||
Db 109 VLALPA 114

RESULT 22
US-08-791-495-5
Sequence 5, Application US/08791495
Patent No. 5811519
GENERAL INFORMATION:
APPLICANT: Leth, Bernard
APPLICANT: Lucas, Sophie
APPLICANT: De Smet, Charles
APPLICANT: Godelaine, Daniele
APPLICANT: Boon-Palleur, Thierry
TITLE OF INVENTION: LL-1 TUMOR SPECIFIC GENES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: US/08/791,495
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Van Amsterdam, John R.
REGISTRATION NUMBER: 40,212
REFERENCE/DOCKET NUMBER: L0461/7005
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 210 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-791-495-5

Query Match 3.5%; Score 6; DB 2; Length 210;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 147 RGGAPR 152

Db 57 RGGAPR 62
|||||
RESULT 23
US-08-840-683-8
Sequence 8, Application US/08840683
Patent No. 5821051
GENERAL INFORMATION:
APPLICANT: Androphy, Elliot J.
APPLICANT: Chen, Jason J.
TITLE OF INVENTION: E6-BINDING PROTEINS
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII (text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/840,683
FILING DATE: 29-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/273,059
FILING DATE: 08-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Louis
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: NEP-003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 220 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
US-08-840-683-8

Query Match 3.5%; Score 6; DB 2; Length 220;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 GPGLSL 11
|||||
Db 68 GPGLSL 73

RESULT 24
US-08-513-974B-315
Sequence 315, Application US/08513974B
Patent No. 6114139
GENERAL INFORMATION:
APPLICANT: Hinuma, Shuji
APPLICANT: Hosoya, Masaki
APPLICANT: Fujii, Ryo
APPLICANT: Ohtaki, Tetsuya
APPLICANT: Fukusumi, Shoji
TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,
PRODUCTION, AND USE THEREOF
NUMBER OF SEQUENCES: 380
CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/513.974B
FILING DATE: 14-SEP-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP95/01599
FILING DATE: 10-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-093989
FILING DATE: 19-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-057186
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-007177
FILING DATE: 20-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-326611
FILING DATE: 28-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-270017
FILING DATE: 02-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-236357
FILING DATE: 30-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-189274
FILING DATE: 11-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-189273
FILING DATE: 11-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S.
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 45753
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 315:
SEQUENCE CHARACTERISTICS:
LENGTH: 223 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-513-974B-315

Query Match 3.5%; Score 6; DB 3; Length 223;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLAUPA 20
|
Db 117 VLAUPA 122

RESULT 25
US-08-513-974B-364
Sequence 364, Application US/08513974B
Patent No. 6114139
GENERAL INFORMATION:
APPLICANT: Hinuma, Shuji
APPLICANT: Hosoya, Masaki
APPLICANT: Fujii, Ryo
APPLICANT: Ohtaki, Tetsuya
APPLICANT: Fukusumi, Shoji
APPLICANT: Ohgi, Kazuhiro
TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,
PRODUCTION, AND USE THEREOF
NUMBER OF SEQUENCES: 380
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/513.974B
FILING DATE: 14-SEP-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP95/01599
FILING DATE: 10-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-093989
FILING DATE: 19-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-057186
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-007177
FILING DATE: 20-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-326611
FILING DATE: 28-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-270017
FILING DATE: 02-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-236357
FILING DATE: 30-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-236356
FILING DATE: 30-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-189274
FILING DATE: 11-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-189273
FILING DATE: 11-AUG-1945
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-189272
FILING DATE: 11-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S.
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 45753
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 364:
SEQUENCE CHARACTERISTICS:
LENGTH: 223 amino acids

TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-513-974B-364

Query Match 3.5%; Score 6; DB 3; Length 223;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLALPA 20
| | | | |
Db 117 VLALPA 122

RESULT 26

US-08-513-974B-368
Sequence 368, Application US/08513974B
Patent No. 6114139

GENERAL INFORMATION:

APPLICANT: Hinuma, Shuji
APPLICANT: Hosoya, Masaki
APPLICANT: Fujii, Ryo
APPLICANT: Ontaki, Tetsuya
APPLICANT: Fukusumi, Shoji
APPLICANT: Ohgi, Kazuhiro

TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,
PRODUCTION, AND USE THEREOF

NUMBER OF SEQUENCES: 380

CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/513,974B
FILING DATE: 14-SEP-1995

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP95/01599
FILING DATE: 10-AUG-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 7-093989
FILING DATE: 19-AUG-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 7-057186
FILING DATE: 16-MAR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 7-007177
FILING DATE: 20-JAN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-326611
FILING DATE: 28-DEC-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-270017
FILING DATE: 02-NOV-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-236357
FILING DATE: 30-SEP-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-236356
FILING DATE: 30-SEP-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-189274
FILING DATE: 11-AUG-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-189273
FILING DATE: 11-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-189272
FILING DATE: 11-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S.
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 45753
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 368:
SEQUENCE CHARACTERISTICS:
LENGTH: 223 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-513-974B-368

Query Match 3.5%; Score 6; DB 3; Length 223;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLALPA 20
| | | | |
Db 117 VLALPA 122

RESULT 27

PCT-US96-12374-2

Sequence 2, Application PC/TUS9612374
GENERAL INFORMATION:

APPLICANT: Northwestern University

TITLE OF INVENTION: Herpes Virus Entry Mediator

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dressler, Goldsmith, Milnamow & Katz, Ltd.
STREET: 180 N. Stetson, Suite 4700

CITY: Chicago

STATE: Illinois

COUNTRY: U.S.A.

ZIP: 60601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/12374

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Northrup, Thomas E.

REGISTRATION NUMBER: 33,268

REFERENCE/DOCKET NUMBER: NOR3446P020PC

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 616-5400

TELEFAX: (312) 616-5460

TELEX: --

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 283 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US96-12374-2

Query Match 3.5%; Score 6; DB 4; Length 283;

Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 CPPGTF 57
Db 165 CPPGTF 170

RESULT 28

US-07-971-092-2
; Sequence 2, Application US/07971092
; Patent No. 5328987
; GENERAL INFORMATION:
; APPLICANT: Maliszewski, Charles R.
; TITLE OF INVENTION: HuIgA Fc Receptor
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex
; STREET: 51 University
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/971.092.
; FILING DATE: 19921104
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia A.
; REGISTRATION NUMBER: 34693
; REFERENCE/DOCKET NUMBER: 2603
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 287 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-971-092-2

Query Match 3.5%; Score 6; DB 1; Length 287;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 LLCLVL 16
Db 8 LLCLVL 13

RESULT 29

5198342-2
; Patent No. 5198342
; APPLICANT: MALISZEWSKI, CHARLES R.
; TITLE OF INVENTION: DNA ENCODING IGA FC RECEPTORS
; NUMBER OF SEQUENCES: 9
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/548,059
; FILING DATE: 05-JUL-1990
; SEQ ID NO: 2:
; LENGTH: 287
5198342-2

Query Match 3.5%; Score 6; DB 5; Length 287;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 LLCLVL 16

Db 8 LLCLVL 13

RESULT 30
US-08-555-722-8
; Sequence 8, Application US/08555722
; Patent No. 5989804
; GENERAL INFORMATION:
; APPLICANT: Androphy, Elliot J.
; APPLICANT: Chen, Jason J.
; TITLE OF INVENTION: E6-BINDING PROTEINS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/555,722
; FILING DATE: 14-NOV-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: NEP-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 317 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-555-722-8

Query Match 3.5%; Score 6; DB 2; Length 317;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 GPGLSL 11
Db 166 GPGLSL 171

RESULT 31

US-08-602-359A-34
; Sequence 34, Application US/08602359A
; Patent No. 5942430
; GENERAL INFORMATION:
; APPLICANT: ROBERTSON, Daniel E.
; APPLICANT: MURPHY, Dennis
; APPLICANT: REID, John
; APPLICANT: MAFFIA, Anthony
; APPLICANT: LINK, Steven
; APPLICANT: SWANSON, Ronald V.
; APPLICANT: WARREN, Patrick V.
; APPLICANT: KOSMOTKA, Anna
; TITLE OF INVENTION: ESTERASES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & RICHARDSON P.C.
; STREET: 4225 EXECUTIVE SQUARE, STE 1400
; CITY: LA JOLLA

STATE: CALIFORNIA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/602,359A
FILING DATE: February 16, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: HALE, LISA A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 09010/010001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-678-5070
TELEFAX: 619-678-5099
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 346 AMINO ACIDS
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-602-359A-34

Query Match 3.5%; Score 6; DB 2; Length 346;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 132 CPPGAG 137
|||||
Db 340 CPPGAG 345

RESULT 32
US-08-513-974B-26
Sequence 26, Application US/08513974B
Patent No. 6114139
GENERAL INFORMATION:
APPLICANT: Hinuma, Shuji
APPLICANT: Hosoya, Masaki
APPLICANT: Fujii, Ryo
APPLICANT: Ohtaki, Tetsuya
APPLICANT: Fukusumi, Shoji
APPLICANT: Ohgi, Kazuhiro
TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,
PRODUCTION, AND USE THEREOF
NUMBER OF SEQUENCES: 380
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patencin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/513,974B
FILING DATE: 14-SEP-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP95/01599
FILING DATE: 10-AUG-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-093989
FILING DATE: 19-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-057186
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-007177
FILING DATE: 20-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-326611
FILING DATE: 28-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-270017
FILING DATE: 02-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-236357
FILING DATE: 30-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-236356
FILING DATE: 30-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-189274
FILING DATE: 11-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-189273
FILING DATE: 11-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-189272
FILING DATE: 11-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S.
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 45753
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 370 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-513-974B-26
Query Match 3.5%; Score 6; DB 3; Length 370;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 15 VLALPA 20
|||||
Db 189 VLALPA 194
RESULT 33
US-08-513-974B-323
Sequence 323, Application US/08513974B
Patent No. 6114139
GENERAL INFORMATION:
APPLICANT: Hinuma, Shuji
APPLICANT: Hosoya, Masaki
APPLICANT: Fujii, Ryo
APPLICANT: Ohtaki, Tetsuya
APPLICANT: Fukusumi, Shoji
APPLICANT: Ohgi, Kazuhiro
TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,
PRODUCTION, AND USE THEREOF
NUMBER OF SEQUENCES: 380
CORRESPONDENCE ADDRESS:
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street

CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/513.974B
FILING DATE: 14-SEP-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-093989
FILING DATE: 10-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-093989
FILING DATE: 19-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-057186
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-007177
FILING DATE: 20-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-326611
FILING DATE: 28-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-270017
FILING DATE: 02-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-236357
FILING DATE: 30-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-236356
FILING DATE: 30-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-189274
FILING DATE: 11-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-189273
FILING DATE: 11-AUG-1945
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6-189272
FILING DATE: 11-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S.
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 45753
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 323:
SEQUENCE CHARACTERISTICS:
LENGTH: 370 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-513-974B-323

Query Match 3.5%; Score 6; DB 3; Length 370;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLALPA 20
|||||
DB 189 VLALPA 194

RESULT 34
US-08-448-722A-4

; Sequence 4, Application US/08448722A
; Patent No. 6072028
; GENERAL INFORMATION:
; APPLICANT: Altieri, Dario C.
; TITLE OF INVENTION: NO. 6072028el Cell Surface Receptor, Antibody
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Office of Patent Counsel, The Scripps
; ADDRESSEE: Research Institute
; STREET: 10550 No. 6072028th Torrey Pines Road, TPC-8
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448.722A
; FILING DATE: 25-MAY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/189,309
; FILING DATE: 28-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: 233.1 Div1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 784-2937
; TELEFAX: (619) 784-9399
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 427 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-448-722A-4

Query Match 3.5%; Score 6; DB 3; Length 427;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 ASCPPG 135
|||||
DB 43 ASCPPG 48

RESULT 35
US-09-012-072-4
; Sequence 4, Application US/09012072
; Patent No. 6060276
; GENERAL INFORMATION:
; APPLICANT: Maslowski, Piotr
; TITLE OF INVENTION: NO. 6060276el Orphan Receptors
; FILE REFERENCE: REG 630
; CURRENT APPLICATION NUMBER: US/09/012,072
; CURRENT FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 4
; LENGTH: 434
; TYPE: PRT
; ORGANISM: HUMAN
US-09-012-072-4

Query Match

3.5%; Score 6; DB 3; Length 434;

Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 PGLSLL 12
Db 230 PGLSLL 235

RESULT 36

US-09-073-569-2
; Sequence 2, Application US/09073569
; Patent No. 6084088
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Grossmann, Angelika
; TITLE OF INVENTION: NOVEL TUMOR ANTIGENS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,569
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sawislak, Deborah A
; REGISTRATION NUMBER: 37,438
; REFERENCE/DOCKET NUMBER: 97-14
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6672
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 437 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-09-073-569-2

Query Match 3.5%; Score 6; DB 3; Length 437;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 PGLSLL 11
Db 92 PGLSLL 97

RESULT 37

5268463-7
; Patent No. 5268463
; APPLICANT: JEFFERSON, RICHARD A.
; TITLE OF INVENTION: PLAAAT PROMOTER a-GLUCURONIDASE GENE
; CONSTRUCT
; NUMBER OF SEQUENCES: 9
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/447,976

; FILING DATE: 08-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 119,102
; FILING DATE: 10-NOV-1987
; APPLICATION NUMBER: 264,586
; FILING DATE: 31-OCT-1988
; SEQ ID NO: 7:
; LENGTH: 457
5268463-7

Query Match 3.5%; Score 6; DB 5; Length 457;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 163 VAGPSL 168
Db 212 VAGPSL 217

RESULT 38

US-08-425-989B-11
; Sequence 11, Application US/08425989B
; Patent No. 5849699
; GENERAL INFORMATION:
; APPLICANT: McClelland, Alan
; APPLICANT: Greve, Jeffrey M.
; TITLE OF INVENTION: Soluble Molecule Related to but
; TITLE OF INVENTION: Distinct from ICAM-1
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bayer Corporation
; STREET: 400 Morgan Lane
; CITY: West Haven
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06516
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" diskette, 1.44 MB storage
; COMPUTER: IBM ThinkPad 760ED
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/425,989B
; FILING DATE: 20-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/156,653
; FILING DATE: 22-NOV-1993
; APPLICATION NUMBER: 08/005,204
; FILING DATE: 15-JAN-1993
; APPLICATION NUMBER: 07/449,356
; FILING DATE: 21-DEC-1989
; APPLICATION NUMBER: 07/445,951
; FILING DATE: 13-DEC-1989
; APPLICATION NUMBER: 07/301,192
; FILING DATE: 24-JAN-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Barbara A. Shimeil
; REGISTRATION NUMBER: 29,862
; REFERENCE/DOCKET NUMBER: MTI 209.2C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203) 812-2786
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 amino acid residues
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: protein
; HYPOTHETICAL: no
; ORIGINAL SOURCE:
; ORGANISM: human

CELL TYPE: epithelial
CELL LINE: HeLa cells
FEATURE:
NAME/KEY: SICAM-1
OTHER INFORMATION: amino acid sequence
OTHER INFORMATION: identical to ICAM-1 protein sequence except
for residue 442, which is Lys rather than
OTHER INFORMATION: Glu, and residues 443-453, which is novel
OTHER INFORMATION: sequence due to alternative splicing
US-08-425-989B-11

Query Match 3.5% Score 6; DB 2; Length 480;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 ALPALL 22
| | | | |
Db 9 ALPALL 14

RESULT 39

PCT-US94-01101-2
Sequence 2, Application PC/TUS9401101
GENERAL INFORMATION:
APPLICANT:
APPLICANT:
APPLICANT:
APPLICANT:
TITLE OF INVENTION: HIGH MOLECULAR WEIGHT B-CELL GROWTH
TITLE OF INVENTION: FACTOR: INTERLEUKIN-14
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
STREET: 600 ATLANTIC AVENUE
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/01101
FILING DATE: FILED HEREWITH
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/005,156
FILING DATE: 15-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: GATES, EDWARD R.
REGISTRATION NUMBER: 31,616
REFERENCE/DOCKET NUMBER: B0819/7000W0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/720-3500
TELEFAX: 617/720-2441
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 498 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: HOMO SAPIENS
FEATURE:
NAME/KEY: Protein
LOCATION: 16..498

PCT-US94-01101-2

Query Match 3.5% Score 6; DB 4; Length 498;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 LVLALP 19
| | | | |
Db 11 LVLALP 16

RESULT 40

US-08-789-078-3
Sequence 3, Application US/08789078
Patent No. 5843885
GENERAL INFORMATION:
APPLICANT: Benedict, Stephen H.
APPLICANT: Slahaan, Teruna
APPLICANT: Chan, Marcia
APPLICANT: Tibbetts, Scott
TITLE OF INVENTION: ICAM-1/LFA-1 PEPTIDES FOR INDUCING
TITLE OF INVENTION: IMMUNE TOLERANCE
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOVEY, WILLIAMS, TIMMONS & COLLINS
STREET: 1101 Walnut St.
CITY: Kansas City
STATE: MO
COUNTRY: USA
ZIP: 64106
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/789,078
FILING DATE: 03-FEB-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/229,513
FILING DATE: 19-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Collins, John M.
REGISTRATION NUMBER: 26262
REFERENCE/DOCKET NUMBER: 22833
TELECOMMUNICATION INFORMATION:
TELEPHONE: (816)474-9050
TELEFAX: 816)474-9057
TELEX: 434-363
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 531 amino acids
TYPE: amino acid
TOPOLOGY: both
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
CELL TYPE: T-cell
CELL LINE: HL-60
FEATURE:
NAME/KEY: Region
LOCATION: 1..25
OTHER INFORMATION: /label= Signal sequence
FEATURE:
NAME/KEY: Region
LOCATION: 41..100
OTHER INFORMATION: /label= Ig1
FEATURE:
NAME/KEY: Region
LOCATION: 128..190
OTHER INFORMATION: /label= Ig2
FEATURE:

NAME/KEY: Region
LOCATION: 230..294
OTHER INFORMATION: /label= Ig3
FEATURE:
NAME/KEY: Region
LOCATION: 325..375
OTHER INFORMATION: /label= Ig4
FEATURE:
NAME/KEY: Region
LOCATION: 413..461
OTHER INFORMATION: /label= Ig5
FEATURE:
NAME/KEY: Duplication
LOCATION: 481..503
OTHER INFORMATION: /label= Trans
OTHER INFORMATION: /note= "Transmembrane domain"
FEATURE:
NAME/KEY: Binding-site
LOCATION: 152..154
OTHER INFORMATION: /label= Attachment
OTHER INFORMATION: /note= "Cell attachment site"
PUBLICATION INFORMATION:
AUTHORS: Pigott,
AUTHORS: Power,
TITLE: ICAM-1 Amino acid sequence (from HL-60)
JOURNAL: The Adhesion Molecule Facts Book
PAGES: 75-75
DATE: 1993
RELEVANT RESIDUES IN SEQ ID NO: 3: FROM 1 TO 531
US-08-789-078-3

Query Match 3.5%; Score 6; DB 2; Length 531;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 ALPALL 22
|||||
Db 9 ALPALL 14

RESULT 41
US-08-752-633-3
Sequence 3, Application US/08752633
Patent No. 5863889
GENERAL INFORMATION:
APPLICANT: Benedict, Stephen H.
APPLICANT: Slahaan, Teruna
APPLICANT: Chan, Marcia
APPLICANT: Tibbetts, Scott
TITLE OF INVENTION: ICAM-1/LFA-1 PEPTIDES FOR INDUCING
TITLE OF INVENTION: IMMUNE TOLERANCE
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESSES:
ADDRESSEE: HOVEY, WILLIAMS, TIMMONS & COLLINS
STREET: 1101 Walnut St.
CITY: Kansas City
STATE: MO
COUNTRY: USA
ZIP: 64106
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/752,633
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Collins, John M.
REGISTRATION NUMBER: 26252
REFERENCE/DOCKET NUMBER: 22833

TELECOMMUNICATION INFORMATION:
TELEPHONE: (816)474-9050
TELEFAX: 816)474-9057
TELEX: 434-363
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 531 amino acids
TYPE: amino acid
TOPOLOGY: both
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
CELL TYPE: T-cell
CELL LINE: HL-60
FEATURE:
NAME/KEY: Region
LOCATION: 1..25
OTHER INFORMATION: /label= Signal
OTHER INFORMATION: /note= "Signal sequence"
FEATURE:
NAME/KEY: Region
LOCATION: 41..100
OTHER INFORMATION: /label= Ig1
FEATURE:
NAME/KEY: Region
LOCATION: 128..190
OTHER INFORMATION: /label= Ig2
FEATURE:
NAME/KEY: Region
LOCATION: 230..294
OTHER INFORMATION: /label= Ig3
FEATURE:
NAME/KEY: Region
LOCATION: 325..375
OTHER INFORMATION: /label= Ig4
FEATURE:
NAME/KEY: Region
LOCATION: 413..461
OTHER INFORMATION: /label= Ig5
FEATURE:
NAME/KEY: Duplication
LOCATION: 481..503
OTHER INFORMATION: /label= Trans
OTHER INFORMATION: /note= "Transmembrane domain"
FEATURE:
NAME/KEY: Binding-site
LOCATION: 152..154
OTHER INFORMATION: /label= Attachment
OTHER INFORMATION: /note= "Cell attachment site"
PUBLICATION INFORMATION:
AUTHORS: Pigott,
AUTHORS: Power,
TITLE: ICAM-1 Amino acid sequence (from HL-60)
JOURNAL: The Adhesion Molecule Facts Book
PAGES: 75-75
DATE: 1993
RELEVANT RESIDUES IN SEQ ID NO: 3: FROM 1 TO 531
US-08-752-633-3

Query Match 3.5%; Score 6; DB 2; Length 531;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 ALPALL 22
|||||
Db 9 ALPALL 14

RESULT 42
PCT-US95-04886-3
Sequence 3, Application PC/TUS9504886
GENERAL INFORMATION:

APPLICANT: Benedict, Stephen H.
APPLICANT: Siahaan, Teruna
APPLICANT: Chan, Marcia
APPLICANT: Tibbetts, Scott
TITLE OF INVENTION: ICAM-1/LFA-1 PEPTIDES FOR INDUCING
IMMUNE TOLERANCE
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOVEY, WILLIAMS, TIMMONS & COLLINS
STREET: 1101 Walnut St.
CITY: Kansas City
STATE: MO
COUNTRY: USA
ZIP: 64106
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04886
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Collins, John M.
REGISTRATION NUMBER: 26262
REFERENCE/DOCKET NUMBER: 22833
TELECOMMUNICATION INFORMATION:
TELEPHONE: (816)474-9050
TELEFAX: (816)474-9057
TELEX: 434-363
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 531 amino acids
TYPE: amino acid
TOPOLOGY: both
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
CELL TYPE: T-cell
CELL LINE: HL-60
FEATURE:
NAME/KEY: Region
LOCATION: 1..25
OTHER INFORMATION: /label= Signal
OTHER INFORMATION: /note= "Signal sequence"
FEATURE:
NAME/KEY: Region
LOCATION: 41..100
OTHER INFORMATION: /label= Ig1
FEATURE:
NAME/KEY: Region
LOCATION: 128..190
OTHER INFORMATION: /label= Ig2
FEATURE:
NAME/KEY: Region
LOCATION: 230..294
OTHER INFORMATION: /label= Ig3
FEATURE:
NAME/KEY: Region
LOCATION: 325..375
OTHER INFORMATION: /label= Ig4
FEATURE:
NAME/KEY: Region
LOCATION: 413..461
OTHER INFORMATION: /label= Ig5
FEATURE:
NAME/KEY: Duplication
LOCATION: 481..503
OTHER INFORMATION: /label= Trans
OTHER INFORMATION: /note= "Transmembrane domain"
FEATURE:
NAME/KEY: Binding-site

LOCATION: 152..154
OTHER INFORMATION: /label= Attachment
OTHER INFORMATION: /note= "Cell attachment site"
PUBLICATION INFORMATION:
AUTHORS: Pigott,
AUTHORS: Power,
TITLE: ICAM-1 Amino acid sequence (from HL-60)
JOURNAL: The Adhesion Molecule Facts Book
PAGES: 75-75
DATE: 1993
RELEVANT RESIDUES IN SEQ ID NO: 3: FROM 1 TO 531
PCT-US95-04886-3

Query Match 3.5%; Score 6; DB 4; Length 531;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 ALPALL 22
Db 9 ALPALL 14

RESULT 43
US-07-618-286-1
Sequence 1, Application US/07618286
Patent No. 5288854
GENERAL INFORMATION:
APPLICANT: Diamond, Michael S
APPLICANT: Staunton, Donald E
APPLICANT: Springer, Timothy A
TITLE OF INVENTION: Functional Derivatives of ICAM-1 Which
TITLE OF INVENTION: Are Substantially Capable of Binding To LFA-1 But Are
TITLE OF INVENTION: Substantially Incapable of Binding to MAC-1 (as amended)
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Ave. Suite 600
CITY: Washington
STATE: D.C.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/618,286
FILING DATE: 19901128
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: MILLMAN, ROBERT A
REGISTRATION NUMBER: 36,217
REFERENCE/DOCKET NUMBER: 1011.0520000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-618-286-1

Query Match 3.5%; Score 6; DB 1; Length 532;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 ALPALL 22
Db 9 ALPALL 14

RESULT 44
US-08-196-003-3
; Sequence 3, Application US/08196003
; Patent No. 5681699
; GENERAL INFORMATION:
; APPLICANT: BEAUDET M.D., AURTHUR L
; APPLICANT: ROTTER M.D., JEROME I
; APPLICANT: TARGAN M.D., STEPHAN R
; APPLICANT: VORA M.D., DEVENDRA
; APPLICANT: YANG M.D., HUIYING
; TITLE OF INVENTION: METHODS OF SCREENING FOR ULCERATIVE
; TITLE OF INVENTION: COLITIS AND CROHN'S DISEASE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PRETTY, SCHROEDER, BRUEGGEMANN & CLARK
; STREET: 444 SOUTH FLOWER STREET, SUITE 2000
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/196,003
; FILING DATE: 11-FEB-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: WHITEFORD ESQ, WENDY A
; REGISTRATION NUMBER: 36,964
; REFERENCE/DOCKET NUMBER: P07 32056
; TELEPHONE: (213) 489-4442
; TELEFAX: (213) 489-4210
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 532 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-196-003-3

Query Match 3.5%; Score 6; DB 1; Length 532;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 ALPALL 22
Db 9 ALPALL 14

RESULT 45
US-08-483-389-118
; Sequence 118, Application US/08483389
; Patent No. 5811517
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Vazeux, Rosemary
; TITLE OF INVENTION: ICAM-RELATED PROTEIN
; NUMBER OF SEQUENCES: 118
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,389
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/102,852
; FILING DATE: 05-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/009,266
; FILING DATE: 22-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/894,061
; FILING DATE: 05-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/889,724
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,689
; FILING DATE: 27-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Suh, Young J.
; REGISTRATION NUMBER: P-41,337
; REFERENCE/DOCKET NUMBER: 27866/32760
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: (312) 474-6600
; INFORMATION FOR SEQ ID NO: 118:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 532 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-483-389-118

Query Match 3.5%; Score 6; DB 2; Length 532;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 ALPALL 22
Db 9 ALPALL 14

RESULT 46
US-08-689-870-12
; Sequence 12, Application US/08689870
; Patent No. 5874233
; GENERAL INFORMATION:
; APPLICANT: Fargan, Stephan R.
; APPLICANT: Vasiliauskas, Eric A.
; APPLICANT: Plevy, Scott E.
; APPLICANT: Yang, Huiying
; APPLICANT: Rotter, Jerome I.
; TITLE OF INVENTION: Methods of Diagnosing a Clinical Subtype
; TITLE OF INVENTION: Of Crohn's Disease with Features of Ulcerative Colitis
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/689,870
FILING DATE: 15-AUG-1996
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-CE 2224
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1599
US-08-689-870-12

Query Match 3.5%; Score 6; DB 2; Length 532;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 ALPALL 22
|||||
DB 9 ALPALL 14

RESULT 47
US-08-933-824-3
Sequence 3, Application US/08933824
Patent No. 6008335
GENERAL INFORMATION:
APPLICANT: BEAUDET M.D., AURTHUR L
APPLICANT: ROTTER M.D., JEROME I
APPLICANT: TARGAN M.D., STEPHAN R
APPLICANT: VORA M.D., DEVENDRA
APPLICANT: YANG M.D., HUIYING
TITLE OF INVENTION: METHODS OF SCREENING FOR ULCERATIVE
TITLE OF INVENTION: COLITIS AND CROHN'S DISEASE
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: PRETTY, SCHROEDER, BRUEGGEMANN & CLARK
STREET: 444 SOUTH FLOWER STREET, SUITE 2000
CITY: LOS ANGELES
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/933,824
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/196,003
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: WHITEFORD ESQ, WENDY A
REGISTRATION NUMBER: 36,964
REFERENCE/DOCKET NUMBER: P07 32056
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-4442
TELEFAX: (213) 489-4210
INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 532 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-933-824-3

Query Match 3.5%; Score 6; DB 3; Length 532;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 ALPALL 22
|||||
DB 9 ALPALL 14

RESULT 48
5284931-2
Patent No. 5284931
APPLICANT: SPRINGER, TIMOTHY A.; ROTHLEIN, ROBERT; MARLIN, STEVEN D.; DUSTIN, MICHAEL L.
TITLE OF INVENTION: INTERCELLULAR ADHESION MOLECULES AND THEIR BINDING LIGANDS
NUMBER OF SEQUENCES: 41
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/515,478
FILING DATE: 27-APR-1990
SEQ ID NO: 2
LENGTH: 532
5284931-2

Query Match 3.5%; Score 6; DB 5; Length 532;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 ALPALL 22
|||||
DB 9 ALPALL 14

RESULT 49
US-08-106-761-2
Sequence 2, Application US/08106761
Patent No. 5445956
GENERAL INFORMATION:
APPLICANT: HAMMOCK, Bruce D.
APPLICANT: GRANT, David F.
APPLICANT: BEETHAM, Jeffrey K.
TITLE OF INVENTION: RECOMBINANT SOLUBLE EPOXIDE HYDROLASE
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND KHOURIE and CREW
STREET: Steuart Street Tower, 20th Floor, One Market
STREET: Plaza
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/106,761
FILING DATE: 19930813
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 2307E-445

TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 554 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-106-761-2

Query Match 3.5%; Score 6; DB 1; Length 554;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLALPA 20
Db 13 VLALPA 18

RESULT 50
US-08-398-008A-2
Sequence 2, Application US/08398008A
Patent No. 5665588
GENERAL INFORMATION:
APPLICANT: Kornbluth, Jacki
TITLE OF INVENTION: DNA Encoding Natural Killer Lytic Associated
TITLE OF INVENTION: Protein
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Gilbreth & Adler, P.C.
STREET: 8011 Candle Lane
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77071
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
COMPUTER: MACINTOSH IICI
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/398,008A
FILING DATE: March 2, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/126,501
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Adler, Dr. Benjamin Aaron
REGISTRATION NUMBER: 35,423
REFERENCE/DOCKET NUMBER: D5705CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (713) 777-2321
TELEFAX: (713) 777-6908
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 587
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: Protein
HYPOTHETICAL: No
ANTI-SENSE: No
US-08-398-008A-2

Query Match 3.5%; Score 6; DB 1; Length 587;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 152 RSGGRR 157
Db 25 RSGGRR 30

Search completed: January 30, 2001, 17:07:10
Job time: 980 sec

THIS PAGE BLANK (USPTO)

1/18

THIS PAGE BLANK (USPTO)